ORBIT

VOLUME 24, NUMBER 2, 1993

A co-publication of the Ontario
Institute for Studies in Education
and Nelson Canada

Standards, Benchmarks, and Ontario Curriculum

A Special Report on Standards and Accountability

ORBIT

A co-publication of the Ontario Institute
for Studies in Education and Nelson Canada

EDITORIAL OFFICES:
ORBIT
252 Bloor Street West
Toronto, Ontario, M5S 1V6

EDITOR: Heather Berkeley
ASSISTANT EDITOR: Donna Hutchins
CONSULTANTS: Ardra Cole, Jane Hill, David E. Hunt,
Peter McLaren, Ruth Scott, and Marielle Simon

GRAPHIC DESIGN: Helmut W. Weyerstrahs

TYPESETTING: Arnheim Typesetting & Design

Printed and bound in Canada
by Webcom
on recycled paper

SUBSCRIPTION RATES:
$32.10 per subscription year (4 issues). Incl. 7% GST

To order: Mail cheque or money order to

ORBIT
P.O. Box 10, Station F
Toronto, Ontario, M4Y 2L4

Phone: (416) 696-2295
Fax: (416) 696-8409

Second Class Mail Registration Number 4511
ISSN 0030-4433 02

NOTES FROM THE DIRECTOR

This issue of *Orbit*, guest-edited by Professor Howard Russell of our Curriculum Department, deals with a topic of great importance to students, parents, and educators — accountability and standards and measuring student academic achievement. Its contributors include OISE faculty and staff, Ministry of Education and Training officials, school board administrators, teachers, and a representative of the business community.

There is a thorough discussion of one aspect of the current debate on assessment — namely, the development of concrete performance standards against which an individual student's efforts can be compared, an approach which diverges from more traditional standardized testing. This approach has been developed over the past five to ten years by the Ministry of Education and Training and by various school boards across the province. Their work is well represented in the issue.

The issue marks the continuation of our co-publication agreement with Nelson Canada.

Arthur Kruger

SPECIAL ISSUE

VOLUME 24, NUMBER 2 1993

STANDARDS, BENCHMARKS, AND ONTARIO CURRICULUM

GUEST EDITOR: *Howard Russell*

1 **Introduction**
Howard Russell

2 **Towards Accountability and Results-Oriented Education**
Developing Provincial Benchmarks
Gail Rappolt

5 **From Curriculum to Outcomes-Based Education**
Provincial Language Standards
Neil Graham

9 **Improving Student Results in Mathematics**
Developing Standards of Achievement
Gérald Hurtubise

11 **Les seuils-repères**
État de la question
Marielle Simon

13 **How Will the Public Respond to Provincial Benchmarks?**
Some Lessons from Public Opinion on Testing
Doug Hart

16 **A Business Perspective on Standards and Accountability**
An Interview with Jim Poole, Ontario Chamber of Commerce

18 **Moving in a New Direction**
An Interview with the Director and the Superintendent of Curriculum, Toronto Board of Education

22 **Demystifying the Goals of Education**
Toronto's Benchmark Program
Sylvia Larter and James Donnelly

25 **Adapting Benchmarks to French Immersion Education**
Working with a Grade 3 Classroom
Sylvia J. Spielman

29 **Surveying the Benchmarks**
An Overview of Programs in Five Ontario Boards of Education
Christine Suurtamm

31 **Assessing Thinking, Understanding, and Problem Solving**
Benchmarks for Grade 2 Geometry
Betty Lou Borho

34 **Linking Opportunities and Achievement**
A Program Evaluation Model for Grade 6 Mathematics
J. Stuart Telfer

36 **"How well are we doing?"**
A Discussion with Bruce Archer, Burle Summers, Ken Travers, and Howard Russell

40 **What Grade for Your School?**
The 1993 Orbit Essay Contest
Elizabeth Campbell

ORBIT

Four
Special
Theme
Issues
for 1993

Only $32.10

ORBIT
Four
Special
Theme
Issues
for 1993
Only $32.10

Standards, Benchmarks, and Ontario Curriculum

Introduction

Howard Russell

Professor of Curriculum, OISE

Benchmarks and standards — which have arrived in Ontario schools during the past decade — constitute a response to the continuing pressure for public accountability in education and are slowly gaining acceptance among teacher groups, trustees, parents, business officials, and governments. Although there are many definitions for benchmarks and for standards, they all fit in the general category of educational indicators of student achievement.

In its use in education, a benchmark is a standard of performance against which students' progress and achievement can be compared. To use more commonplace terminology, the Ontario Ministry of Education and Training has adopted the term "standards" to replace "benchmarks," and to indicate continuity and avoid further confusion, now uses the term "standards (benchmarks)." In this issue of *Orbit*, we use the terms standards and benchmarks interchangeably except where boards of education specifically use the term "benchmark" only.

Benchmarks and standards focus on curriculum substance in an open manner that permits, and perhaps encourages, public debate about what students are taught and what they learn as well as what general directions education should take. The more traditional indicators have focussed on marks and general scores which cannot be related to changes needed in curricula or programs.

This issue of *Orbit* begins with a section on the Ontario Ministry of Education and Training's initiatives. The lead article by Gail Rappolt focusses attention on the Ontario standards. She provides the Ministry's background discussion, principles, and process in the development of provincial standards. The principles upon which they were developed constitute an especially important starting point. The articles by Neil Graham and Gerry Hurtubise tell us what students are to learn and teachers are to teach in language and mathematics, respectively. These clear accounts of outcomes will be used also in debates about what should be taught in our schools. The article by Marielle Simon outlines the Ministry's benchmarks initiative in French and describes some local initiatives in French-language school boards.

The second section indicates some of the perceptions and expectations of the public and of the business community. Doug Hart presents public opinion poll information which provides readers with the best available evidence on what the public wants. The public wants to know what their own particular children have learned and they want the teachers of those individual children to have a role in determining what is taught and how students are evaluated. They also want to know what impact our provincial programs have on student learning and they want to be able to make comparisons over time and across provincial and national boundaries. Our authors in this issue of *Orbit* have addressed these questions.

The voice of the business community is presented through an interview with Jim Poole, who represents the Ontario Chamber of Commerce on various committees dealing with education. His insights into the issue of public accountability in education are based on careful analysis of a vast array of up-to-date information, and as a consequence they may be both revealing and generally acceptable to educators.

The third section of this issue deals with the Toronto Board of Education's benchmarks Program. It begins with an interview with the Director of Education, Joan Green, and the Superintendent of Curriculum, Marilyn Sullivan. We have given special attention to the Toronto program because it has a long history behind it and it has given other school systems, both here and abroad, an opportunity to see how a diverse curriculum in a cosmopolitan city can be monitored effectively. Ms. Green's comments on the acceptance of common benchmarks across racial and ethnic lines are both helpful and insightful. The Donnelly and Larter article and the Spielman article spell out some of the details that show the nature of Toronto benchmarks.

The fourth section explores some other initiatives at the school and board levels. The first article is based on a paper Christine Suurtamm prepared for the 1992 Annual Meeting of the Ontario Educational Research Council. It covers five boards that have established benchmark programs and it outlines some of their essential features. Betty Lou Borho brings some clarity and realism to what may be labelled "real mathematics" in the primary grades. She labels her benchmarks the "s" list. The "s" signifies "school" and it helps to distinguish among objectives on the basis of where the learning takes place. Stuart Telfer shows how benchmark information is used at the school board level. His benchmarks are in fact test items in the usual sense, and he presents some novel ideas for connecting curriculum change to benchmarks.

The issue concludes with a discussion on standards, benchmarks, and curriculum with Bruce Archer, Burle Summers, Ken Travers, and myself. The view of benchmarks and standards from Washington and the National Science Foundation, as expressed by Ken Travers, indicates the basic similarity between the accountability issues we face in Ontario and those faced in other jurisdictions.

In a sense, we have defined benchmarks and standards by giving examples. The rest of the story will be the array of uses and reporting processes that emerge over time. That part of the story has yet to come. We are very much interested in our readers' comments on the substance of this issue and we encourage you to write to us care of the Editor of *Orbit* at OISE.

The Ontario Ministry of Education and Training's Standards (Benchmarks) Project

Towards Accountability and Results-Oriented Education

Developing Provincial Standards

Gail Rappolt

Field Co-ordinator
Provincial Benchmarks Project,
Ontario Ministry of Education
and Training

It has been stated that the 1990s will be the decade of accountability. Nowhere will that be more apparent than in education. A recent article in *The Economist* identified the quality of a county's education system as the way countries will be measured in the 90s. It went on to compare student achievement in the 90s to GNP in the 70s or 80s as a measure of a country's international prominence. The combination of two phenomena — the world-wide focus on quality, results, and accountability and the emphasis on educating all of society to a much higher level and for a quite different world — has put tremendous pressure on education systems, schools, and educators.

The response at all levels has been to turn much more attention to the actual outcomes we want, to the standards we expect and the student results we are getting, and to how to measure and report those results. Many school boards in Ontario began work on such projects in the late 1980s or early 1990s, the most widely known of these being the Toronto Benchmarks Project. In the fall of 1991, the Ministry of Education began a standards project as well, at that time known as the Provincial Benchmarks Project. The intention was to draw on the work already done in England, Scotland, some provinces in Canada, and some Ontario school boards to develop a comprehensive set of publicly agreed upon standards that would be part of an overall assessment or evaluation framework.

The first task was to determine the principles or beliefs on which an assessment system should be based. Over time the following set of principles emerged:

Standards referenced

The definition of the standard should come before the assessment. In other words, we are going to agree on a predetermined standard and then test against it. We are not merely going to test and rank results against each other.

Improvement focussed

The primary purpose of assessment should be to improve results — for the student, the school, and the jurisdiction. Although that sounds obvious, many of the large assessments currently conducted do not provide either the type or the quality of information that leads to improvement; they merely provide scores to rank performers against each other or against norm or average.

Fair to all

Assessment systems must be sensitive to class, race, culture, and gender differences and to disabilities; they must not penalize any group in subtle or covert ways.

Valid

Assessment systems must assess the outcomes that are worth assessing and use appropriate tools to do so. Implicit in this statement is the recognition that multiple-choice tests, indeed paper and pencil tests alone, cannot assess all the outcomes of significance. Embedded in the principle of validity is reliability: to be reliable, the assessments need to yield consistent results.

Relevant

Assessment systems must be used and reported in the context of other relevant information. For example, information on student results taken out of context says little about the quality of instruction or value for money.

Collaborative

Parents and the community should be involved in setting the standards. Teachers should be involved in designing the assessment systems.

Understandable

Procedures and results should be understandable to teachers and parents as well as to researchers and administrators. Results should be public and reported against pre-determined standards. Reporting results against other schools or jurisdictions leads to what is called "high-stakes testing" and undermines several of the other principles (standards referenced, improvement focussed, valid, reliable).

Refined regularly

The pace of change in all aspects of today's world has altered traditional views about using the same test over decades for purposes of reliability. Assessment systems and standards themselves should be subject to continuous review and improvement.

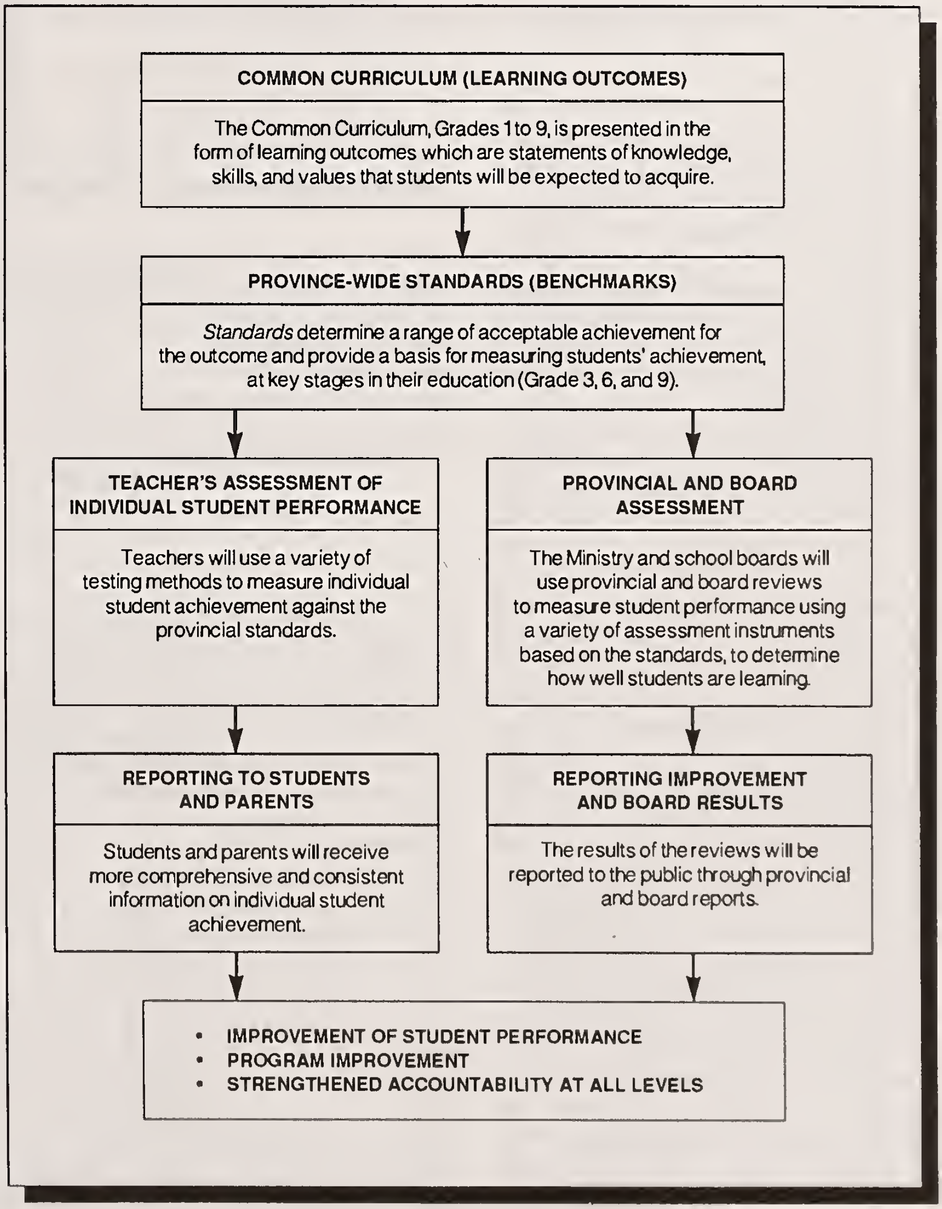
The second task of the provincial project was to develop a framework in which standards and assessment would fit (see Figure 1). It was clear that the framework had to reflect the results or outcomes focus of the 90s and to differentiate between outcomes and standards. There were several major shifts in thinking that had to occur for both educators and the public.

The first was the shift from what we called an "intentions-based" curriculum (in Ontario represented by the Formative Years and OS:IS), where the curriculum is stated in terms of objectives and the teacher is the person responsible for the action. Most objectives at the Ministry level have been written in this way — for example, "students will have opportunities to ..." or "students will experience ...". Here the onus is on the teacher to act, perform, and deliver the result. In a shift to an "outcomes-based" system, the student is now accountable for the outcome — for example, "students will demonstrate awareness of the four cueing systems when reading." Implicit in this student performance or achievement is the teacher's role in ensuring that these skills are taught and learned; however, the student must take the actual action and demonstrate his or her achievement.

The second shift was to separate the general result that is desired (the outcome) from the level of performance that is expected (the standard). Whether or not there are two separate concepts here is still hotly debated in some circles, though researchers like Grant Wiggins make a very clear differentiation (see his 1991 article "Standards, Not Standardization"). It was the contention of the Provincial Benchmarks Project that there needed to be two quite separate concepts articulated. We further suggested that for now, as we make the transition from an intentions- and subject-based curriculum to an outcomes-based and cross-disciplinary curriculum, we need to keep subject-based standards as a support for both teachers and parents.

The third shift was to demonstrate that we could use the same standards of performance for measuring student achievement in the classroom as we would use to monitor programs in schools, school boards, or the province. Previous assessment frameworks had kept these two types of assessment separate partly to avoid high-stakes testing and partly because the large-scale achievement tests had not tended to measure complex concepts, higher-level thinking skills, or attitudes. What made this idea so difficult to accept was that people did not differentiate between standards and standardized (tests). In addition, we had to demonstrate that these "standards" were acceptable against international standards. Responding to a demand for international comparability while attempting to measure the complex skills and performance that are demanded in the Common Curriculum will continue to be a very challenging task. In their recently released book *Overdue Assignment*, Jennifer Lewington and Graham Orpwood use some superb examples and analogies in the chapter called "Measuring Suc-

Figure 1 *Accountability Flowchart*



COMMON CURRICULUM (LEARNING OUTCOMES)
PROVINCE-WIDE STANDARDS (BENCHMARKS)
TEACHER'S ASSESSMENT OF INDIVIDUAL STUDENT PERFORMANCE
PROVINCIAL AND BOARD ASSESSMENT
REPORTING TO STUDENTS AND PARENTS
REPORTING IMPROVEMENT AND BOARD RESULTS

The Common Curriculum, Grades 1 to 9, is presented in the form of learning outcomes which are statements of knowledge, skills, and values that students will be expected to acquire.

Standards determine a range of acceptable achievement for the outcome and provide a basis for measuring students' achievement, at key stages in their education (Grade 3, 6, and 9).

Teachers will use a variety of testing methods to measure individual student achievement against the provincial standards.

The Ministry and school boards will use provincial and board reviews to measure student performance using a variety of assessment instruments based on the standards, to determine how well students are learning.

Students and parents will receive more comprehensive and consistent information on individual student achievement.

The results of the reviews will be reported to the public through provincial and board reports.

- IMPROVEMENT OF STUDENT PERFORMANCE
- PROGRAM IMPROVEMENT
- STRENGTHENED ACCOUNTABILITY AT ALL LEVELS

cess" to explain terminology and to clear up misunderstandings.

The fourth task was to differentiate between "standards," "standardization," and "standardized tests." Over the past ten months, the major media have grappled with the notion of "standards versus standardized." The issue has been further complicated by the fact that the term "standardized," as it refers to testing, has been broadened by researchers in some circles to mean any large-scale test. Teachers tend to think of standardized tests as norm-referenced, paper and pencil, group, multiple-choice tests. Teachers question their validity and their relevance, and Ontario students don't get much practice doing them. Lack of skill in doing these tests could be one of the reasons why Ontario students have not performed well on such tests. Parents and others tend to refer to any test or exam that a large number of students participate in as a standardized test and some poll results suggest that parents want more "standardized tests." In the Benchmarks Project, we are talking about *standards*! In fact, we are talking about high, worthwhile, and consistent standards for all, even if some (or many) cannot yet reach them (Wiggins, 1991). This has been a very difficult shift for people to make.

Figure 2 *Five Known Methods of Developing Standards*

1. A *number*, a norm, often an average. This number is used as a standard against which to measure performance over time, or as a basis for ranking (e.g., I.Q. test scores).
2. A *criterion*. A number reference or set of criteria against which performance is measured (e.g., words per minute in typing).
3. *Tacit knowledge*. The knowledge of the expert or connoisseur. These standards are resident in the head of the expert, but are not available to others (e.g., olympic diving).
4. *Verbal description*. A statement makes the standard explicit, public and accessible so the standard can be known, met, challenged (e.g., job descriptions).
5. *Key examples*. Examples of the performance or behaviour are shown (visual or text) to help people recognize the standard when they see it. They are not explicit standards, they *represent* standards (e.g., videos of performance).

The reasons for the difficulty are obvious and they lead to the next shift in thinking. What is a standard and how are we defining a standard? We have used the work of Grant Wiggins (cited above) to define a standard and the work of D. R. Sadler (1987) in the *Oxford Review of Education* as the basis for determining how standards would be set. What we learned as we worked with Wiggins's notion across the province was that very few people could conceive of a standard as anything but a number or a rank. Sadler identifies five ways of setting standards and we have tried to give some examples of each in Figure 2. Once we had done the research and identified the outcomes expected by futurists, professionals, and teachers, we developed the initial provincial benchmarks or standards by eliciting the tacit knowledge from expert teachers (type 3). These standards were then written down (verbal statements, type 4) and indicators were added to make them more explicit. Often to determine the indicators or evidence (criteria, type 2), we had to go to key examples of student work (type 5) or to norm-referenced test results (type 1). It was our intent to embrace all the different types of standards in the project so that it would not be viewed as narrow or only serving certain purposes. Sadler's work provided a way for us to be consistent with our own principles.

The next shift in thinking was to have teachers, parents, and the public come to a common understanding about the terms test and assessment. Again, the chapter on accountability in *Overdue Assignment* provides clarity about the appropriate uses and the differences. A test, as the word is traditionally used, measures performance under controlled conditions at one point in time. Assessment implies a broad range of instruments and methods used to used to ensure that the desired standard has been met and that the achievement or performance is consistent over time. Performance assessment usually refers to the measurement of complex rather than simple tasks.

Finally, in the framework, the shift in the way we think about reporting had to be dealt with. There is a great deal of controversy about how results should be reported. There is not a clear differentiation in the mind of educators or the public of different types of accountability — fiscal accountability, performance management of staff, or quality of student results. Information about context must be available before decisions or judgments can be made on the basis of data. If we accept the principles on which the project is based, then data can only be reported for purposes of improvement and only against the standards. (See Figure 3.)

The data are not relevant unless they are in context. Given the current world

Figure 3 *From Data to Decisions and Actions*

	TYPES	
STEPS	**STUDENT** (Teacher Level)	**PROGRAM** (School, Board, Provincial Level)
Assessment	Collection of a wide range of data about an individual student's achievement	Collection of a wide range of data about overall achievement in the jurisdiction
Evaluation	Reporting to • student • parent about individual achievement as referenced to a standard	Reporting to • teachers • principals • boards • public about achievement of the jurisdiction as referenced to a standard
Decisions	• remediation • enrichment • placement • promotion	• change policy • provide staff development • realign funding

and economic context described in the opening paragraph of this article, it will be very hard to maintain these principles at the provincial, school board, or school level. If we don't, however, we will end up spending a great deal of money testing against standards where the data, when reported, tell us very little about actual achievement. This will happen because schools will focus on ensuring a high rank in the reporting of results instead of focussing on improving programs at the individual child, school, or system level.

In the end, if the principles and framework are to result in improved student achievement, then we must move from outcomes to standards, from assessment to data, to decisions, and to action.

This description of the development process is a retrospective one. The process as it evolved was messy, slow, often two steps forward and one back. We were breaking new ground, working at tasks and in ways that we believe are consistent with education systems around the world as they struggle to prepare their students (and their teachers) for the 21st century. The real test will be the product; the draft provincial standards which have been developed are described in two other articles in this issue. They will be available to schools and across the province in the fall of 1993 as interim standards in Grades 3, 6, and 9 mathematics and language. They will be reviewed by teachers, parents, and the public, and ultimately it will be the Royal Commission on Learning that recommends the accountability measures that will remain or move into place in Ontario in the future.

REFERENCES

Lewington, J., & Orpwood, G. (1993). *Overdue assignment*. Rexdale, ON: John Wiley & Sons.

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, 13(2).

Wiggins, G. (1991, February). Standards, not standardization: Evoking quality student work. *Educational Leadership*, pp. 18–25.

From Curriculum to Outcomes-Based Education

Provincial Language Standards

Neil Graham

Learning Assessment Branch, Ontario Ministry of Education and Training

Throughout the English-speaking world, the last ten years have seen a dramatic shift from curriculum-based education to outcomes-based education. In Ontario, this move has had a powerful impact on education as teachers and administrators at all levels attempt to develop standards and assessment methods which focus on outcomes and at the same time acknowledge the complexity of learning and of assessment. Those who were involved in any way in the 1991/92 Provincial Review of Grade 12 Writing will be well aware of this complexity. Designing and administering the assessment activity, assessing students' writing, and setting the provincial standards against which to judge their writing performance all involved complex but necessary procedures that ensured that the link between learning and assessment was maintained.

Over the past three years, the Ministry of Education and Training has been working on this shift in emphasis from curriculum to outcomes-based education on two fronts: developing provincial standards (benchmarks) in mathematics and language, and developing a Common Curriculum which reorganizes curriculum delivery for Grades 3, 6, and 9 and articulates specific outcomes for the end of these divisions.

Because the development of mathematics standards began a year earlier than language standards, they are at different stages of development. Language standards have been under development for the past year and have just now reached the stage where draft material can be presented to teachers, administrators, and the public for response. Over the next 18 months, both the mathematics and language standards will be treated as interim standards, subject to modification and refinement. The draft language benchmarks (see Figure 3) will be available in schools in the fall. In the meantime, this article provides an overview of the development to date and a first-draft sample.

Creating a Framework for Provincial Language Standards

During the past year, the Ministry of Education and Training has had two committees working on the development of provincial language standards for Grades 3, 6, and 9. A steering committee representing a variety of educational groups has provided leadership and direction in the articulation of guidelines upon which to base the development of the standards, and a writing committee of classroom teachers and consultants has developed the provincial language standards from the guidelines and framework established by the steering committee.

The following material provides an overview of the conceptual framework and some sample material the committees have developed for establishing and using provincial language standards to assess student performance.

Guidelines for Creating Provincial Language Standards for Grades 3, 6, and 9

After considerable deliberation and review of existing approaches to standards-setting from various provincial, national, and international jurisdictions, the steering committee agreed on the following guidelines for developing the provincial language standards for the Province of Ontario:

1. The language standards would be based on the language outcomes described for Grades 3, 6, and 9 in the Common Curriculum. These outcomes are described in terms of knowledge, skills, and values that are consistent with the developmental stages of language acquisition.

2. One set of language standards would be developed for all students, including ESL/ESD students and special education students. The standards would accommodate the present Ontario practice of age-appropriate placement of these students.

3. Although reading, writing, and speaking and listening activities are integrated and should be assessed in an integrated manner, for purposes of assessing performance, separate provincial language standards would be developed for each of reading, writing, and speaking and listening.

4. For each of Grades 3, 6, and 9, the provincial language standard for each language area would be described in terms of an expected range of performance, in effect establishing an all-inclusive provincial performance scale or a scale for each language area and each identified grade level. The scale would describe six levels of performance, with the standard — that is, the *expected* range of performance for that grade — being levels three, four, and five. Performance at levels one and two would be considered to not yet be at the expected performance, and level six performance to be beyond the range of expected performance for that grade. Specific assessment procedures would require the development of specific scales developed from pertinent information found in the scales which describe the total performance. (See Figures 1 and 2.)

Figure 1 *The Link Between Outcomes, Standards, Assessment, and Reporting*

DETERMINING OUTCOMES	Outcomes state general results expected.	• The Common Curriculum (Grades 1-9) determines the knowledge, skills and values expected for all.
SETTING STANDARDS	Standard state the expected range of student achievement for the outcomes.	• The province will provide clear standards in language and mathematics.
ASSESSING PERFORMANCE	Performance will be assessed against the provincial standards in language and math (Grades 3, 6, and 9).	• Classroom teachers will have common provincial standards as a basis for evaluating and reporting student achievement. • School boards and the province will use the same standards for board-wide and provincial testing.
REPORTING RESULTS	Results will be reported in relation to the provincial standards in language and math (Grades 3, 6, and 9).	• Classroom teachers will, where appropriate, report student achievement in relation to the standards. • School boards and the province will report school-wide, school board and provincial performance in relation to the standards.

5. The scales would demonstrate gradations in performance within the six levels for each of the three identified grades and demonstrate horizontal coherence in expected performance across the three grades. However, the scales across the grade level should not be developed in an overlapping manner; for example, level five in Grade 3 should not be seen as a level two performance in Grade 6.

Developing the Provincial Language Standards For Grades 3, 6, and 9

Using the various methods by which standards are developed (see Gail Rappolt's article for a description of those methods), the writing committee has articulated six levels of performance for each language area as follows.

Each level is identified by a word that describes the quality of the overall performance. This is followed by a general descriptor indicating what the overall performance in the specific language

Figure 2 *Provincial Language Standards Framework*

LEVEL 6 Superior Performance	Students performing beyond expected range of standards set for grade.
LEVEL 5 Proficient Performance	**PROVINCIAL STANDARDS** Students performing within the *expected* range set for grade.
LEVEL 4 Capable Performance	
LEVEL 3 Adequate Performance	
LEVEL 2 Limited Performance	Students not yet performing within expected range set for grade
LEVEL 1 Rudimentary Performance	

Figure 3 *Draft Language Standards (draft June 4, 1993)*

PROVINCIAL STANDARDS - GRADE 9 WRITING AND REPRESENTING

Level 6 - Superior Writing Performance

The student is an independent and self-confident writer who has internalized and adapted the various stages of the writing process to the point where they are applied automatically. The writing exhibits a strong voice and a command of a wide range of forms, genres, language conventions, and stylistic devices. The writing is thoughtful, clear, imaginative, and often innovative.

Indicators

Characteristics:
- The content is compelling. The focus is complex and is supported by a wealth of original and perspicuous detail.
- The organization is expert. There is some surprising subtlety in the arrangement.
- The writing evidences an impressive command of the resources of language: a wide vocabulary, a full range of sentence patterns, control over stylistic devices.
- The writing evidence complete control over the conventions of the form in which it is operating.

The Process:
- The writer monitors his or her own growth as a writer; sets long range goals for improvement and tracks own progress toward those goals.
- The writer can access all the features of desk top publishing and graphics programs to integrate art, graphics and text in a polished piece of work.

Attitude:
- Strongly values and enjoys writing. Appreciates the full range of purposes of writing; to discover, shape and communicate personal meaning; to enter into and contribute to discourse communities; offer to access information and ideas across time and space.

Level 5 - Proficient Writing Performance

The student is an independent and self-confident writer who has largely internalized the various stages of the writing process. The writing exhibits a clear voice and a command of a range of forms and genres, language conventions and some stylistic devices. The writing is thoughtful, clear, and imaginative.

Indicators

Characteristics:
- Has a fairly complex focus and a wealth of relevant and convincing supporting detail.
- Has a complex hierarchical organization that is sustained throughout.
- Is a written with full considerable command of the resources the language; it can vary the tone, mood, style to suit the purpose.
- Has a fine grained appreciation of the characteristics of the genre on which it is operating.

The Process:
- Understands the writing process to include long range planning as well as a recursive approach to individual pieces.
- Understands and can draw on a variety of strategies for generating and improving text.
- Can access mechanical style and grammar analysis programs

Attitude:
- Values and enjoys writing appreciates role on shaping personal and cultural identity; enjoys and contributes to community of writers in the classroom.

Level 4 - Capable Writing Performance

The student is largely an independent and self-confident writer, depending on the writing task, and is generally skilful in applying the stages of the writing process. The writing exhibits some sense of voice and a control of a range of forms, genres, language conventions, and some stylistic devices are present. The writing is generally thoughtful, clear, and, at times, imaginative.

Indicators

Characteristics:
- The content has clear focus with relevant supporting detail.
- The material is discernably organized in an orderly progression of sections with explicit transitions.
- It is written in correct and continuous prose that complies with the conventions of spelling, punctuation and grammar.
- It demonstrates a familiarity with the conventions of the form or genre in which it is operating.

The Process:
- Understands the processes of planning, drafting, revising.
- Seeks out advice and feedback at all stages of the process.
- Can access regular features of word processing.

Attitude:
- Values the place of writing in school and society.

Continued on next page

Figure 3 *Continued*

Level 3 - Adequate Writing Performance

The student has some self-confidence as a writer and is generally skilful in applying most of the various stages of the writing process. The writing exhibits a sense of voice at times and a control of basic forms, genres, and language conventions, and some attempt at using stylistic devices is evident. The writing is generally clear.

Indicators

Characteristics:
- Has a sustained organization throughout, using simple ordering and listing arrangements.
- Has discernable form and is working within the recognizable characteristic conventions of a genre.
- Is written in continuous prose.
- The writer exercises considerable control over vocabulary, spelling, and conventional punctuation.

The Process:
- Has an organized system for storing work in progress and saving completed work.
- Understands revising to include improving substantive content and organization as well as surface features.
- Regularly revises working in progress.

Attitude:
- Is beginning to appreciate writing as a learning tool.

Level 2 - Limited Writing Performance

The student is a developing writer who has some self-confidence with certain aspects of writing and is able to apply some of the stages of the writing process with guidance. The writing exhibits capability with some of the basic forms, genres, and language conventions. The writing demonstrates clarity at times.

Indicators

Characteristics:
- There is an attempt at organization: effective opening and closing sections and some attempts to order the material in between around a main idea.
- Words are strung together in intelligible phrases and sentences but the conventions of spelling, marking sentence boundaries are not evident.
- There is some attempt at form, but only the broadest characteristics of the form are evident.
- The writer revises work independently or collaboratively to correct surface features: spelling and punctuation.

Attitude:
- The writer is beginning to gain personal satisfaction from writing.

Level 1 - Rudimentary Writing Performance

The student is an emerging writer who can apply some of the stages of the writing process with guidance. The writing exhibits a rudimentary grasp of some of the basic forms, genres, and application of a few of the basic language conventions.

Indicators

Characteristics:
- Sentence fragments
- Minimal information with shifting focus.
- Organization limited to primitive opening and closing statements.
- Has limited words and structures available for expressing meaning; controls everyday vocabulary words but has difficulty stringing them together.
- Has a limited and undifferentiated sense of form or genre and audience.
- Prefers oral to written expression.
- Unconventional spelling of frequently used vocabulary.

The Process:
- Will revise work with support, direction, and encouragement.
- Revises by recopying to improve legibility.

Attitude:
- Needs to be encouraged and motivated to write.
- Prefers writing that relates to personal experience.

area looks like, and that descriptor is followed by a set of indicators providing detailed description of various elements of the performance. The general scale describing provincial standards for Grade 9 Writing (see Figure 3) demonstrates this pattern. (Keep in mind that at this stage, the scales are written in professional language for use by teachers.)

Using the Provincial Language Standards

Teachers and jurisdictions wishing to assess student performance against provincial standards will use these generic performance descriptors and select and/or adapt the indicators that apply. For example, if their assessment activity were broad enough to include all the indicators identified in any one language area — for instance, assessing student writing through writing folder assessment — they would be able to use the general scale as is. If, however, the assessment activity were narrower — a single piece of writing or reading passage or an oral presentation — the descriptors and/or indicators would need to be narrowed to what is being measured by the assessment.

The Ministry will use the general provincial scales and the identified provincial standards for reading and for writing to assess students' level of performance on the Grade 9 Reading and Writing test this fall. Since the test asks students to respond to a variety of reading passages and to write a variety of compositions over a period of time, the general descriptors and most of the indictors will apply.

Follow-up and Response

The committees will complete work on these provincial standards over the summer and in the early fall schools will receive a complete set of draft materials for reading, writing, and speaking and listening for Grades 3, 6, and 9. Teachers and interested groups are asked to examine these interim standards over the next 18 months and respond to them.

Anyone wishing to respond to the overview material presented in this article is asked to contact the Learning Assessment Branch, Ministry of Education and Training, 18th Floor Mowat Block, 900 Bay Street, Toronto M7A 1L2; (416) 327-9071.

Improving Student Results in Mathematics

Developing Standards of Achievement

Gérald Hurtubise

Eastern Ontario Regional Office,
Ontario Ministry of Education
and Training

In response to public pressure to deliver improved student results in mathematics, the Minister of Education requested, in the fall of 1991, that provincial standards for mathematics (for Grades 3, 6, and 9) be the first priority of the Provincial Benchmarks Project. In February, 1992, increased impetus for the math project came with the release of the results from the International Assessment of Educational Progress and the Provincial Review results for Grades 8, 10, and 12.

The purpose of developing mathematics benchmarks is the improvement of student results in mathematics. Standards will be used to identify areas where instruction can be improved, to diagnose shortcomings, to determine necessary improvements, and to provide means for the enrichment and enhancement of mathematics programs in both English- and French-language schools across the province. They are to be standards for student achievement and assessment as well as standards for the monitoring of programs.

The development and production of these standards are the result of close collaboration among mathematics teachers, Ministry of Education and Training officials, and representatives of the three mathematics associations (two English and one French) in Ontario.[1]

In researching and developing the mathematics standards, the research team looked at curriculum and assessment materials from school boards in Ontario and other provinces in Canada. Work from other English and French speaking countries was carefully analysed. Where possible, translations and summaries of directions in Asian and other European countries were reviewed. It is important to note, however, that although material from other countries, particularly those whose values and school culture are radically different from Ontario's, can yield valuable information, it cannot provide simple recipes to follow.

The document "Provincial Mathematics Benchmarks" provides a framework for standards in mathematics for Grades 3, 6, and 9. The material in the document is organized in two ways: by grade and by components of mathematics as defined by Ontario mathematics educators. These components have been grouped into the following six *strands* in the document:

1 Inquiry and Problem Solving

2 Number Sense and Numeration

3 Spatial Sense and Geometry

4 Patterning and Algebra

5 Data Management and Probability

6 Estimation and Measurement

For each strand, for each grade, there are one or two statements of expected learning outcomes (tied to the outcomes in the Common Curriculum), statements of the standards or benchmarks, sample evidence to clarify these statements, and sample assessment activities that give key examples to indicate whether the standard had been met. (See Figure 1; sample assessment activities have not been included in this article.)

The benchmark statements themselves indicate a provincial standard for *all* students, followed by additional achievement for *most* students. A further level of achievement, which *some* students will attain by the end of the grade, is also described. This three-part structure is intended to accommodate the views of teachers and community groups who recommended that the provincial standard of achievement should not be set impossibly high, but that neither should only a minimum level of achievement be accepted.

The standards document does not discuss or recommend methods of instruction to be used in the classroom. However, the performances described in the benchmarks indirectly indicate the experiences or practice students need to have in order to meet the standards. Figure 2 outlines the fundamental principles of mathematics education presented by the Ontario Mathematics Co-ordinators Association. The principles are consistent with the beliefs and teaching standards expressed by the National Council of Teachers of Mathematics in their standards documents. (They also seem consistent with the views expressed by Ken Travers of the National Science Foundation in this issue of *Orbit.*)

Figure 1 *Primary Division, Strand 2 (draft)*

CRITICAL LEARNING OUTCOMES
Understand the fundamental operations of addition, subtraction, multiplication and division and their inter-relationships.

BENCHMARKS	SAMPLE EVIDENCE
All students will: given a real-life context and appropriate resources, demonstrate the fundamental operations accurately.	— demonstrate an understanding of addition and subtraction with concrete materials. — know basic facts without the use of concrete materials or calculators as defined below: Addition: up to 10 + 10 Subtraction: up to 20 - 10 Multiplication: up to 5 x 5. — demonstrate the addition and subtraction of two 2- and 3-digit numbers using concrete materials (with and without regrouping). — demonstrate understanding of multiplication and its relationship to addition using concrete materials. — using concrete materials or representations, demonstrate multiplication and division. — identify the operation required to solve a simple problem. — —
In addition, most students will: given a real-life context and appropriate resources, perform the fundamental operations accurately and explain.	— demonstrate the addition and subtraction of 2- and 3-digit numbers using paper and pencil procedures (with and without regrouping). — demonstrate an understanding of division and its relationship to subtraction using concrete materials. — know basic facts, without the use of concrete materials or calculators, as defined below: Division: up to 25 ÷ 5. — identify and use the specific operations of addition and subtraction in a problem solving situation. — create and solve problems using the operations of addition and subtraction. — use the relationship between related operations to verify results. — —
In addition, some students will: given a real-life context, perform the fundamental operations accurately without using resources, and explain and demonstrate the computation.	— know basic facts as defined below: Multiplication: up to 10 x 5 Division: up to 50 ÷ 5. — create and solve problems which demonstrate the operations (+,-,x,÷) in a given situation. — add and subtract two numbers with more than 3 digits (with or without regrouping). — —

Figure 2 *Fundamental Principles of Mathematics Education Grades 1 to 9*

Throughout the process of instruction and student assessment for mathematics, it is assumed that:

- all learning will be within a developmental framework
- technology, problem solving, estimation strategies will be integrated throughout
- students will be working with a wide variety of manipulative materials
- students will be looking for patterns on which to make generalizations and predictions
- the teaching strategies are activity based, collaborative and co-operative
- the learning tasks are authentic and integrated
- conclusions will be justified verbally or in writing

The standards also reflect the mathematics component of the learning outcomes embedded in the Common Curriculum in the Math, Science, and Technology core program area. The learning outcomes in the Common Curriculum describe achievement and performance that cross the traditional lines of subjects; the benchmarks ensure that the basic knowledge, skills, and beliefs necessary for competence in mathematics are achieved.

The process of setting publicly agreed upon standards is far from complete. The draft document was developed to provide a basis for discussion in order to refine and validate the standards it contains. It was circulated for comment as widely as possible during the fall of 1992, accompanied by a validation response form. This form asked general questions about the framework, structure, and approach of the document and also asked for very specific comment on the levels of student achievement expected in each of the standards.

During the winter of 1992/93 over 525 responses to the document were reviewed. The results of the analysis are now being turned over to a team of teachers who will rewrite the standards based on public input. In very broad terms, the results were as follows:

- In general, there was very strong support for the development of clear and publicly agreed upon standards.
- There was, however, a general lack of understanding and agreement about what a *standard* actually is and for what purposes it should be used.
- Although there was general agreement that the standards should identify a range of performance, rather than a single level, the use of the terms "all students," "most students," and "some students" was generally described as unacceptable.
- There were opposing opinions expressed about the appropriate weight given to geometry, data management, patterning and algebra, and complex problem solving in relation to "the basics" of numeracy and measurement.
- Many educators expressed the view that the standards, as written, were not currently attainable by our students; those outside education tended to rate the standards as too low overall.

There is also a problem of language. How can mathematics educators use the language of mathematics to state high and consistent standards in the discipline and still meet the need for plain, clearly understandable language? That is a dilemma yet to be resolved.

During the summer of 1993, revised standards that take all the concerns and diverse opinions into consideration will be prepared. Some of the changes will be straightforward; some will require judgment and therefore not please everyone. The current work being done internationally on mathematics performance standards and the work that has been done over the last year in developing the language standards for Ontario have provided new insights and have led to a slightly revised structure for the mathematics standards.

The Royal Commission on Learning, formed in early May 1993, will review the broad issues of educational accountability over the next 18 months. The draft provincial mathematics standards will be interim standards during the term of the Royal Commission and will be available in the schools in early fall, 1993. They will provide standards against which to measure individual student's performance and a basis for discussion between parents and teachers. As well, they will be the basis for discussion in the broader community — discussion that will have to answer two questions:

> "Do these standards measure the important mathematical outcomes we want students to learn?"
>
> "Are these mathematics standards high enough?"

If you have comments or questions about the Ontario Mathematics Standards or if you would like a copy of the draft document in French or in English, please contact Gérald Hurtubise, Eastern Ontario Regional Office, Ministry of Education and Training, 4th Floor, 1580 Merivale Road, Nepean, Ontario, K2G 4B5; (613) 225-9210, extension 190.

NOTE

1. The co-operation and assistance of the teachers, researchers, mathematics co-ordinators, and members of the Ontario Association for Mathematics Education who contributed to the development of the mathematics standards have been very much appreciated by Ministry staff. Also TVO's December 4, 1992, program "Responding to the Provincial Mathematics Benchmarks" has been a useful staff development tool.

Les seuils-repères

État de la question

Marielle Simon

Directrice Centre régional d'Ottawa
Institut d'études pédagogiques de l'Ontario

Le concept de seuils-repères tel que l'entend le ministère de l'Éducation de l'Ontario est apparu lors de la refonte en éducation, à la suite de l'étude des résultats des élèves ontariens aux enquêtes provinciales, nationales et internationales des apprentissages. Les seuils-repères se veulent des «points de référence» permettant d'assurer cohérence, continuité et uniformité dans l'enseignement et de préciser le degré ou la qualité du rendement des élèves de la province par rapport aux attentes des milieux scolaire et d'affaires.

Le présent article fait état de la question des seuils-repères provinciaux en mathématiques pour la fin de la 3e, 6e et 9e année et décrit certaines initiatives régionales en Ontario français. Pour ce faire, nous avons d'abord consulté Monsieur Gérald Hurtubise, agent d'éducation au bureau régional d'Ottawa du Ministère d'Éducation de l'Ontario et responsable du dossier en question. Nous avons aussi visionné l'émission Pédagogie 2000 télédiffusée en novembre 1992, qui mettait en vedette Monsieur Hurtubise et Madame Constance Legentil, coordonnatrice des programmes du Conseil des écoles séparées d'Essex. Nous avons demandé à des représentants/es des conseils scolaires et des sections de langue française de la région de l'est de l'Ontario de bien vouloir partager avec nous les démarches entreprises dans le domaine.

Seuils-repères provinciaux

L'établissement de seuils-repères sert à atteindre quatre buts. Ils permettent de rassurer le public ontarien de la qualité des programmes-cadres du Ministère de l'Éducation de l'Ontario; ils guident l'élaboration des enquêtes provinciales; ils amènent l'enseignant/e à situer l'élève dans son cheminement et fournissent aux enseignants/es une variété d'outils et de pistes permettant d'ajuster les pratiques pédagogiques.

Le ministère de l'Éducation s'est d'abord penché sur l'élaboration des seuils-repères provinciaux en mathématiques. Le processus d'élaboration et de validation comprenait trois étapes:

a) Mise sur pied d'équipes d'élaboration des seuils-repères

Un groupe d'enseignants/es de la

ORBIT

Did you see our last issue?

Violence in the Schools/ Schooling in Violence

Guest Editor W. Gordon West, Associate Professor, Department of Sociology, OISE, and Centre of Criminology, University of Toronto

featuring

- Black Students and School Violence
- Gender and Violence in the Schoolyard
- Homophobic Schools: Factories of Violence
- Social Class, Streaming, and Violence
- On Teaching Violence and the Media
- Conflict Resolution: Problems and Strategies at the Secondary Level
- The Peaceful School: An Elementary School Perspective
- Bullying at School: Pervasive and Persistent
- Preventing Violence Towards Children: The Role of Education

Plus Dialogues on the Nature and Extent of the Problem with School Administrators, Teacher Federation Representatives, Social Service Workers, the Police, Sociologists, and School Children

Copies still available at $5.35 (GST included)

Or Four Special Theme Issues for 1993 for only $32.10 (including GST)

Mail your order to Orbit, P.O. Box 10, Station F, Toronto M4Y 2L4
Fax (416) 696-8409

ORDER NOW!

région de Windsor a élaboré les seuils-repères provinciaux en mathématiques 3e année. Un groupe d'enseignants/es de la région de North Bay et Timmins a rédigé ceux de mathématiques 6e année. Un groupe d'enseignants de la région d'Ottawa s'est occupé de développer les seuils-repères mathématiques 9e. Une fois que les indicateurs furent identifiés de façon parallèle dans les deux communautés linguistiques, il y a eu fusionnement des deux études.

b) Étude critique des seuils-répères des Conseils scolaires de l'Ontario, des programmes-cadres et des programmes d'études des autres provinces canadiennes et de plusieurs pays occidentaux et orientaux

Les équipes comprenaient des enseignants/es, des membres d'associations dans les domaines d'éducation et du travail ainsi que des personnes-ressource compétentes en mathématiques.

Bien que les seuil-répères se définissaient à partir des pratiques d'enseignement précédentes et actuelles en mathématiques, on a tenu à élargir les cadres consultatifs pour faire en sorte que les seuils-répères soient à la hauteur des attentes locales et universelles.

c) Processus de consultation et de validation auprès des divers groupes intéressés

Une fois les critères formulés, le ministère de l'Éducation de l'Ontario distribuait une ébauche à tous les groupes intéressés pour fins d'examen critique. Cet exercice permettait d'orienter les politiques ministérielles et de réviser les seuils-repères de sorte qu'ils soient complets et pertinents.

Les seuils-repères provinciaux ont trois volets. Une première série de seuils sert à identifier ce que *tous* les élèves d'un niveau scolaire devraient être capable de faire à la fin de l'année scolaire. Une deuxième vise la *plupart* des élèves tandis que pour la troisième série, on s'attend à ce que *quelques* élèves seulement manifestent les comportements décrits. Cette marge de manoeuvre respecte les habiletés de tous les élèves.

On prévoit plusieurs retombées de l'utilisation des seuils-repères. Au niveau provincial, ils mèneront à un enseignement des mathématiques plus cohérent et uniforme. Ils renseigneront la clientèle de ce qui est attendu des élèves au niveau des connaissances, des compétences et des attitudes à certains moments fixes. Finalement, les conseils scolaires en province pourront s'inspirer des seuils-repères provinciaux lors de l'élaboration de stratégies locales d'enseignement et d'évaluation. On espère que l'utilisation des seuils amènera un changement dans la façon de penser vis-à-vis l'enseignement des mathématiques. Il ne suffira plus de suivre fidèlement et uniquement un manuel scolaire ou un programme d'étude mais de planifier l'enseignement de sorte que tous les élèves possèdent les savoirs, les savoirs-faire et les savoirs-être désirés.

On se penche actuellement sur l'élaboration de seuils-repères provinciaux en communication écrite et orale.

Seuil-repères locaux

Plusieurs conseils scolaires ou sections de langue française en Ontario ont déjà élaboré des seuils-repères locaux dans certaines matières. Dans l'est de l'Ontario certains conseils ou sections ont établi entre autres, des seuils en chimie cycle supérieur, en biologie cycle supérieur, en mathématiques 9e année, en anglais/English 10e année niveau général et en français cycle moyen.

Dans un conseil en particulier, par exemple, une équipe de spécialistes a cerné d'abord les dimensions essentielles des programmes d'études. On a ensuite identifié les critères propres à chaque dimension et définit la performance minimale attendue à partir de l'analyse critique des travaux des élèves et des épreuves d'évaluation. On a aussi fourni des exemples de questions ou d'instruments pour chacun des seuils. Ces seuils ont été mis à l'essai pendant un an.

Dans un autre conseil la participation à des ateliers en évaluation des apprentissages a amené les conseillers/ères pédagogiques à concevoir les seuils-repères en termes d'indices ou d'indicateurs de réussite.

Dans la plupart des conseils scolaires cependant, on préfère attendre la mise en oeuvre officielle des seuils-repères provinciaux et du programme d'études commun de la 1e à la 9e année. On élaborera ou adaptera alors les programmes d'études ainsi que les stratégies d'enseignement et d'évaluation à la lumière de ces documents.

How Will the Public Respond to Provincial Benchmarks?

Some Lessons from Public Opinion on Testing

Doug Hart

Senior Research Officer,
OISE Survey of Educational Issues

An important goal of the Provincial Standards (Benchmarks) program (and local Board programs) is to raise public confidence in the school system. Whether this succeeds will, in large part, depend on the extent to which benchmarks are seen by the public as increasing the accountability of the school system. At present, there is virtually no information on how the public at large (as opposed to organized interest groups) will react to the provincial approach. In the absence of direct data, however, it is worth re-examining public opinion on issues of testing for clues to likely public reaction to benchmarks. Most of the following is based on opinion data from the OISE Surveys of Educational Issues, conducted every two years and last administered in September, 1992.[1]

In predicting public response to the benchmarks initiative, four lessons can be drawn from the record of public opinion on testing issues:

1. Benchmarks will be readily accepted insofar as they are associated in the public mind with province-wide testing.
2. Benchmarks will encounter widespread opposition if they are seen as supplanting or seriously compromising local autonomy in student assessment and, in particular, the role of teachers.
3. Only a minority of the public will hold opinions which are "consistent" — by the standards of organized interest groups — across issues of benchmarks, role of teachers, control over curriculum, province-wide testing, and so on. Thus, while at the level of public debate support for local control of curriculum typically goes together with opposition to province-wide tests, opinions on these issues may be only weakly correlated among the mass public.
4. Parents and non-parents will make similar judgments about benchmarks, suggesting that long-standing orientations to assessment, rather than direct interest and experience with the school system, are pivotal in opinion formation on assessment issues.

That benchmarks can be readily sold to the public if packaged as, or with, province-wide testing will undoubtedly be seen as obvious by most readers. It is hardly news that polls have consistently found strong majority support for testing. Two recent surveys confirm the pattern. An Angus Reid poll released in early February found that 77 percent of Ontario respondents (and 76 percent nationally) would support "... standardized testing in our educational system" (*Toronto Star*, 1993, p. A16). And in September, 1992, 73 percent of respondents to the OISE Survey of Educational Issues agreed that "province-wide tests should be used to assess the individual performance of elementary school students." This represented a sharp increase from 1990, when the figure was 59 percent. The 1992 OISE Survey also found that support for testing was particularly high among those who thought that the most useful information to parents was how their child's skills stacked up against official standards or benchmarks. Note that "benchmarks" were defined generally in the survey as official standards of what skills should have been learned. The questionnaire items did not refer to the provincial benchmark initiative explicitly. See Figure 2 for the exact wording of the items.

There is, however, more to public views on province-wide testing than these simple, and much publicized, results. While an overwhelming majority support province-wide testing, only a very small minority want such tests to be the most important means of assessing students. In 1990, only 9 percent of respondents to the OISE Survey wanted provincial tests to be most important in judging the academic progress of *elementary* school students. (This question was not asked in 1992.) Similarly small minorities in the 1979, 1982, and 1988 OISE Surveys had supported giving priority to provincial tests in assessing students "in the higher grades" (Livingstone et al., 1989, p. 21). In 1988, these views on the *best* means of assessing students co-existed with majority support for the use of province-wide tests. A survey conducted early in 1991 by the Institute for Social Research (ISR), York University, found that only 13 percent thought students' progress should be judged by province-wide tests alone (ISR, 1991, p. 4). This option is admittedly more extreme than the position taken by most proponents of testing.

Thus, the second lesson to be learned from opinion research on testing is that support for province-wide testing does not translate into an identification of accountability with meeting externally set standards. Issues of local input and influence on student assessment are also important. Both the OISE survey and ISR questions juxtaposed province-wide tests with local options; in the ISR question, the alternative was "parents and teachers" as "... best to judge a student's progress." Twenty-nine percent selected this option; 52 percent opted for both province-wide tests and parents and teachers. The options for the OISE question are shown in Figure 1.

Just over half of respondents support a wholly locally based option as the best way to assess elementary students' academic progress. The proportions were similar in 1979, 1982, and 1988 when the question was asked about students in the higher grades. One-third think that the best way to assess progress involves joint local/provincial responsibility, while only a small minority would hand the job to the province alone. While the majority of the public are wedded to the test format as the best means of student assessment, local teacher-made tests alone or in conjunction with province-wide tests are seen as key.

Figure 1 *"Which one of the following do you think is the best way of judging the academic progress of students in elementary school?" (Asked in 1990)*

	All respondents %	Parents %	Non-parents %
The individual teacher should have sole responsibility for making assessments	20	20	20
Assessments should be based on tests set jointly by teachers in the school	34	37	31
Provincial tests should be used to supplement teacher-made tests	35	34	35
Provincial tests should be most important	9	7	11
Don't know/ not stated	3	3	3

Figure 2 *Support for Use of Province-wide Tests at the Elementary Level by Views on Most Useful Information to Parents of Elementary School Students*

	Most useful information to parents % agreeing with province-wide tests			
	All respondents	Parents	Non-parents	Number of respondents*
How their child's skills compare to how they were doing earlier	64	63	65	491
How their child's skills compare to those of other students	76	73	77	160
How their child's skills compare to official standards or benchmarks	85	85	88	288
Don't know/ not stated	75	54	82	62

* The sample included 396 parents of elementary and/or high school students.

A majority of the public are strongly supportive of the role of local teachers in student assessment yet they also favour the use of province-wide tests. These positions do not usually go together in the polarized debates over testing. Thus, the third lesson to be learned from public opinion on testing is that the public often holds positions on issues which do not go together "logically," at least according to those at the centre of public debates.

In 1984, the OISE Survey first asked whether province-wide tests should be used to assess the individual performance of high school students. The question was immediately preceded by one on who should have the most influence on what is taught in local schools. Respondents were almost equally split over whether local groups (school board, local teachers, parents) or the provincial or federal government should have the most influence over the curriculum. Those who chose one of the local options were less likely to support use of province wide-tests than others, but the difference was narrow: 63 versus 73 percent. In other words, a substantial majority of those who favoured a dominant local influence on curriculum also favoured province-wide tests, which are usually regarded in public debates as an instrument of centralized control over what is taught in local schools.

The 1992 OISE Survey included a more general question on whether the involvement of various groups in the operation of local schools should increase, decrease, or stay the same. Many more respondents favoured the increased involvement of parents (84 percent) and teachers (69 percent) than the increased involvement of the Ministry of Education (41 percent) and school board administrators (39 percent). In other words, support was highest for those in immediate contact with students. Views on what should happen to the involvement of different groups in the operation of local schools had no bearing on opinion regarding province-wide testing. Those who wanted more involvement of teachers and parents in the operation of local schools were no less likely than others to support province-wide testing; those who favoured more Ministry involvement were no more likely to indicate support for testing.

The 1992 Survey also asked respondents to select the single most useful information to parents regarding how their child is doing in school. The options are shown in Figure 2. Those who selected the option which focussed on a child's progress in relation to his or her own past performance were least likely to support province-wide testing. However, whether respondents held child-centred, norm-oriented, or benchmark-oriented priorities, a majority supported province-wide tests. This is the case for both parents and non-parents. Those who selected child-centred assessment comprised over 40 percent of those supporting province-wide testing.

That public views are frequently, even typically, "inconsistent" when judged by the standards of opinion-leaders has been a common finding of opinion research. From one perspective, this reflects the "inattentiveness" of mass publics to views advanced by opinion-leaders. From another, inconsistencies are the counterpart of overly rigid positions which have evolved in an increasingly polarized debate. In either case, these patterns counsel caution in interpreting the reaction of public opinion to provincial standards and benchmarks. It is important to note that new testing initiatives will likely be evaluated against what appear to be quite stable and complex general orientations to student assessment.

This leads to our final point. It is unlikely that the views of organized interest groups on benchmarks will simply be adopted by large segments of the public. Yet general, abstracted orientations to assessment rather than concrete interests and recent direct experience with the school are likely to be pivotal in opinion formation. This is suggested by the fact that despite their much more direct contact and intensive involvement with the school system, the views of parents of elementary and high school students differ little from those of non-parents on assessment issues. Parents are as likely to favour use of province-wide tests as non-parents. As noted in Tables 1 and 2, the distribution of views on the best method of assessing students' academic progress and the most useful information to parents about their child's progress are similar for parents and non-parents. "Don't know" and non-response rates on assessment questions are low for both groups. That neither direct experience nor "attentiveness" to organized interest groups will play a key role in attitude formation suggests that much of the public will attempt to put whatever "new wine" comes with benchmarks into "old bottles."

NOTE

1. The nine OISE surveys have all been published and distributed by OISE Press. They are as follows: D. W. Livingstone. *Public Attitudes Toward Education in Ontario 1978*; D. W. Livingstone and D. Hart. *Public Attitudes Toward Education in Ontario 1979, 1980*. D. W. Livingstone, D. Hart, and L. D. McLean. *Public Attitudes Toward Education in Ontario 1982*. D. W. Livingstone, D. Hart, and L. E. Davie. *Public Attitudes Toward Education in Ontario 1984, 1986, 1988, 1990*, and *1992*. For more detailed information about the design of the OISE survey series, see especially D. W. Livingstone, D. Hart, and L. E. Davie (1985). *Public Attitudes Toward Education in Ontario 1984: Fifth OISE Survey*. Toronto: OISE Press.

REFERENCES

Toronto Star. (1993, February 5). [Angus Reid poll results], p. A16.

Institute for Social Research. (1991, April 23). [Media release]. Toronto: Author.

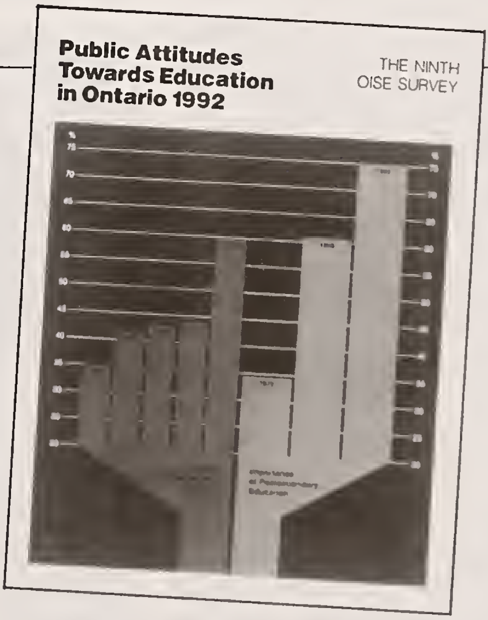
Livingstone, D. W., Hart, D., & Davie, L. (1989). Public attitudes toward education: 1988. *Orbit*, 20(1).

The Ninth OISE Survey

Includes Current Trends and Opinions on:

- Educational Funding
- Elementary School Curriculum and Testing
- High School Destreaming
- Equity Issues
- Postsecondary Access
- Community College Programs
- Adult and Continuing Education
- Politics of Education

PLUS
views on other current issues, including educational roles of television and libraries, and possible social service roles for schools.



0-7744-0396-9/$7.00
(+ $2.00 handling, + 7% GST)

To order *"Public Attitudes"* telephone, write or fax:
OISE Press, Scholarly Book Services Inc.,77 Mowat Avenue, Suite 403 Toronto, Ontario M6K 3E3
Tel: (416) 533-5490
Fax: (416) 533-5652

A Business Perspective on Standards and Accountability

An Interview with Jim Poole, Ontario Chamber of Commerce

Jim Poole

Member of the Education Policy Committee of the Ontario Chamber of Commerce and Representative on the Assessment Advisory Committee of the Learning Assessment Branch, Ontario Ministry of Education and Training

Interviewed by Howard Russell

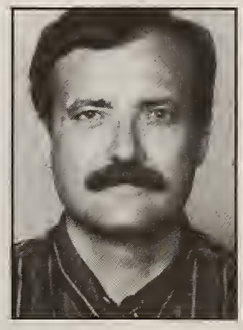
Russell What do you think the Ontario Chamber of Commerce wants from education authorities in order to satisfy the need for accountability?

Poole The Chamber of Commerce wants an education system for Ontario which is accountable, and the best way to show this is through benchmarks, that is, performance standards and achievement levels, which are reported publicly. The largest problem is the lack of information, both from a reporting sense, as well as a perceived lack of information generated within the system to provide feedback to our teachers, educators, and administrators to help improve the system. The business community wants to receive graduates to help compete globally against nations that presently score higher than Canada on international tests and it supports, as a first step towards accountability, regular formal assessment of student progress and publication of overall results in core subjects through standardized provincial systems. These should be put in place at the earliest stages of elementary education.

To get going on that, we need a better definition of curriculum and core subjects starting in the primary years and then we need to measure progress towards mastery of the curriculum.

Russell Do you see evidence of progress towards the satisfaction of this need?

Poole Where I see the evidence is in the direction taken by the Ministry of Education and Training through the provincial benchmarks — now called "standards" — program, which has been running on a number of subjects for at least the last five years. It provides information for educators in the system after Grades 3, 6, and 9. That is excellent feedback to give to education system managers so they can help improve the system. I believe it's important to decide on a task, look at the results, and implement the learning into the system, and then try to continuously improve the process. Our committee has been supporting the Ministry on this. I believe that the difficulty is for parents and business — they feel it's not far enough, not fast enough, not detailed enough, especially in the primary years. Yet we're encouraged by the number of professionals in the system who want information, accountability, and freedom to help improve the system.



— I believe too many groups have their fingers in the education pie.

– Jim Poole

Russell Is the benchmarks or standards thrust distinguished from others in the past with respect to the clarity of goals and the outcomes of education? I'm thinking now of standardized tests with the label "math" or "literacy" associated with school marks.

Poole The business community, parents, and all other stakeholders in the system will know more if the results are communicated properly through all channels. I believe one of the difficulties is that the system goes off and takes the information into itself and doesn't communicate it outside, and we're left with a vacuum of knowledge that is currently raising people's concerns. It's obvious there are barriers to progress, because results from system change haven't been achieved to date, and I believe that's because too many groups have their fingers in the education pie. This leads to a lack of trust among participants in the process. What is needed is to set benchmarks or standards and get information flowing so the system can use the information to improve. In my view, there are a lot of opinions floating around, but not enough facts. With information, the level of trust will certainly build, reduce barriers, and help speed change. The Chamber of Commerce wants to help in this regard by building consensus between labour, education, and business in a nonpartisan setting. We believe the politicization of the education process, which changes with each government and Minister, is in fact a barrier to progress. Politicians are involved and are using it to further the political process. We think education should rise above that and be a longer term direction of the province.

Russell Does the business community use benchmarks or indicators? Your experience in the business community may tell us something about how trust can be built.

Poole The education community, I think, goes through a similar process. You plan the curriculum. You have a longer term need. You introduce the curriculum. You check to see how students are achieving and then you make changes in whatever the delivery process is to improve results. I think that process is very similar to the process in business to measure progress and track performance towards a goal. The business community learned that it exists only to serve the customer, and successful companies do everything to help their front-line employees delight their clients. Business tries to get the right information to these workers so they understand their achievement and are accountable for continuous progress towards goals that they help to set. There's ownership of these goals. Many businesses seek feedback or survey their customers or consumers of services so they can "benchmark" progress towards improvement.

The business community wants to receive graduates to help compete globally against nations that presently score higher than Canada on international tests

How does this apply to the education system? I believe educators should understand the needs of students and graduates and the customers of the education system — business, universities, community colleges, even drop-outs of the system; these are just dissatisfied consumers. This will help educators gain knowledge about the performance of the education system.

Russell In education, we begin with some lack of trust between educators and their clients and we may end up eroding the trust that exists when the benchmark information — the very kind of information educators need to improve the service they provide — is distributed in a way that is misleading but newsworthy. Does this kind of erosion take place in the business community?

Poole The issue of not communicating fully is, in my mind, the base issue behind the erosion of trust throughout the system. Without real data and information, the partners in the education system become suspicious of each other's motives. You get many that are trying very hard to do good work but that counts for nothing if it is not communicated properly. I think that points to an alliance between all of the stakeholders in the system to share common learnings, if I can put it in general terms here.

Russell What are the most important features of a benchmarks program or what are the features that must be preserved at all cost?

Poole The most important features of the benchmark program are, first, that it communicates progress to the public, including business, and that's how we raise confidence in the system. The second is that it provides feedback to the system so professionals can improve the education progress. The third is that it gives feedback to individual students and parents against a set of meaningful standards, by grade, so that parents understand what their children are capable of. I think those are three things that have to be preserved. If something had to be eliminated due to cost, I would suggest individual testing for every student could be eliminated and we could use statistical sampling techniques to cut costs with the same accuracy of data.

Russell Do you see a connection between educational data and workplace data when we move in education towards benchmark-level data as opposed to macro-level data?

Poole When we collect data, the best possible indicator of achievement is something that is knife-edged — either you did it or you didn't. We work very hard to get indicators that give that progress. Getting data to track progress towards a plan and to evolve an organization or process towards a desired outcome is, in fact, the kind of measure we in business apply to our processes, and I see the use of similar statistical data as being supportive in both education and business communities. The effort to gain an understanding of the system before the action was taken, and then after, and determining how well the curriculum passed through to the student is exactly the kind of statistical evidence you require in order to manage the education system; you need the feedback for teachers to improve the process.

Russell Do you have any advice to teachers who are tempted to begin work with benchmarks?

Poole My advice is to start. No benchmark system is going to be perfect the first time. So, getting started on it — understanding the system, improving the benchmarking system with feedback to the originators — is the best place to start. You end up with a much improved system with that kind of experience. To work towards achievement of real data through benchmarking, or setting standards, and keeping records of overall achievement and then to work to improve the accuracy and achievement of this process is what I think educators should be driving towards. When business uses benchmarks — and we do benchmark processes — it's not to compare numbers but to first benchmark the process in order to understand it and to lead to a better result.

Russell Let's talk a bit about that. Do we have enough clarity about what we mean by benchmark and is it going to be intelligible to the business community?

Poole When we talk about benchmarking, we talk about going into a company that has a similar function and looking at a process — billing to the customer, for example — at how it works, what are the benefits, what's good and bad, and then we turn around and look at our process, put it up side by side, and see whether there are areas of improvement in process; then we measure outcomes towards that. I don't believe the theory is any different in the education community. I believe that what we call a "benchmark" is the measurement of outcome of the process.

Russell Any concluding comments?

Poole The Ontario Chamber of Commerce believes very strongly that the benchmark or standards program in the Ministry of Education and Training is a step in the right direction, that information flow, both internal and external, is essential to manage the education system, and that can be achieved through information generated by benchmarks. The Ontario Chamber of Commerce wants to help build consensus towards action between education, labour, and business.

The Toronto Board of Education's Benchmark Program

Moving in a New Direction

An Interview with the Director and the Superintendent of Curriculum, Toronto Board of Education

Joan Green
Toronto Board of Education

Marilyn Sullivan
Superintendent — Curriculum,
Toronto Board of Education

Interviewed by Howard Russell

Howard Russell

Russell I'd like to begin by asking you to describe the Toronto Board benchmark program.

Green We define our benchmarks as strategies or tools to help us establish standards of achievement for students in our board. We see benchmarks as providing examples of the kinds of performances that lead to academic success as well as system-wide reference points to assist our teachers in student evaluation. All of us involved in education these days believe that we need to make both what we're doing and the results more readily understandable by all concerned, particularly by the parents and the caregivers of the children in our schools. It's no longer enough for us to understand what children are achieving. It is absolutely paramount that parents and the children themselves have a clear notion of how they're doing relative to others. Our benchmarks are helping us move in that direction.

Russell I see the Toronto Board Benchmarks as marking a fundamental shift in the nature of accountability data. You've taken a lot of heat for moving in that new direction.

Green Yes, we've paid a big price both in terms of actual dollars and in human resources. We believe it's a price well invested. We think we've developed an eclectic array of strategies that capture many of children's capacities and reflect their ability to perform in a wide range of areas, not just in a narrow band. I think we all know that as we look at the demands of a global economy and the demands of living in the 21st century, there is going to be no single track for performance. It will be a requirement for young people to take their knowledge from one context to another with ease and to operate in a problem-solving mode. We believe very strongly that what we're doing and how we're assessing that must reflect those realities for our kids as they leave our schools. So there's no point in having a set of evaluative measures that test only pen and paper skills, or "fill-in-the-blank" skills, when we know the world is far more complex than that.

Sullivan I think it's important for people to know that we conducted genunine research to develop our benchmarks. We found out what the range of performances was on each individual task and then determined the percentages of children that were operating at each level. The descriptions that are then gained for each one of those levels of performance become criteria for teachers to judge future performances of students in their classrooms.

Russell What progress have you made over this last ten-year period? Do you see we're near the end of it, or just beginning, or how would you describe it?

Green We see benchmarks as not simply tests. We see them as ways to gain insight into student performance and achievement. And we also see them as providing a framework within which teachers can reflect on their own programs and develop a window on their activity and their methodology. They're really intended, in the end, to serve as appropriate models for assessment in every classroom. In the past five years, we've tried to make benchmarks understandable to all of our leaders and to individuals who are responsible for the professional growth of staff in every school setting. We've actually completed one cycle of very rigorous in-service for all of our consultative staff, for principals and vice-principals, and we've made the implementation of benchmarks site-based. Yet it's a long-term process and I suppose our current frustration is that we're not sure if we've managed to help teachers translate what they've learned into language that is accessible to parents. Our biggest challenge now is to make sure we get past the jargon associated with any professional language and make our benchmarks clearly understandable to parents.

Russell It sounds like you've had to grapple with a communication barrier, or at times a conceptual barrier. I'd like you to give us a little more insight into dealing with this problem.

Green We've had to grapple with getting a common understanding among our teachers. It wasn't there to begin with and we had to introduce them to a very complex interconnected set of evaluation strategies. So first we had an education problem or challenge with our own teachers. We had to create all kinds of opportunities for teachers to talk, to come to that understanding.

Sullivan In other words, in the first phase we tried to ensure that all our administrators, curriculum staff, and teachers really understood the philosophy, principles, and practices related to benchmarks. The second phase really is more related to evaluation and to ensuring that we move from benchmarks as good models of classroom practice to using benchmarks in the evaluation of daily work. Then there's the reporting to parents phase. We're conducting a pilot in mathematics and language in two different families of schools using a portfolio system that allows teachers to collect samples of student work and to track student progress based on benchmarks. The feedback from those two pilots will allow us to move into a system-wide implementation at some point early next year. Our early feedback from teachers, parents, and principals is that they finally see how the total program can be used to help them understand what children are learning in the classroom and how well their students are progressing in relation to system standards.

Russell How about the community at large? You've said that teachers understand the direction you are taking but what about others in the community?

Green That's the point I was trying to make a few moments ago, that we have a translation task before us and particularly with parents of children who are disadvantaged in one way or another. They believe that their children are not having the same advantages, outcomes, or results from the school experience as other children. So they're a bit impatient with the detail and the obscure language that sometimes accompanies professional talk. We are trying very hard to make things clear. In the last little while we developed a resource sheet for teachers, "Benchmarks and the Basics," which talks about what it is that parents define as basics. Some people define the basics with a sort of myopic nostalgia. We are trying to help them see that the basics are given in the Toronto school system, but that what's required in the future will be more than basic skills and, therefore, the evaluation tools we use have to assess more than what would be defined as the old basic skills.

Russell Do minority groups in Toronto have a distinct position on standards and benchmarks?

Green The leaders of our different parent groups believe that their children should be performing at the same level, with the same range of opportunities, as other children have traditionally. So they don't want remedial programs. They want their children to be taking full advantage of all of our programs. They also want their children's evaluation to reflect that experience. We have an organization in Toronto called the Organization of Parents of Black Children. It is probably one of the most helpful groups we've dealt with and also one of the most challenging. They worked with our staff to produce a report, now several years old, on the education of black students, and one of its key tenets and recommendations is that we focus on ensuring that children of all groups, and particularly

Toronto benchmarks are not the static school-bound definitions that we had in the past but they are real, dynamic, vital and allow and encourage lively discussion between parent and teacher.

— Joan Green

black children, have the same opportunities and that we demonstrate the same results. Their fear is that young people of colour could be clustered in one setting where the standards would not be the same as in other settings.

Russell What about other groups in the community, business for example? What is their response to the Toronto benchmarks?

Green Right now I'm working closely with the Metro Toronto Learning Partnership which consists of directors of education (three of us serve on the steering committee) and CEOs of major corporations in Ontario, such as Xerox, Bell Ontario, Northern Telecom, Peat Marwick Thorne. We're working to establish an effective partnership to intervene in school settings and help with school improvement. Now, you can imagine the politics of that. Many people fear a business take-over, or an imposition of a business agenda, and therefore an erosion of the pure definition of education. But one of the areas we've been working on very hard is the whole question of what standards mean and how business views standards. We've used the Conference Board of Canada's statement of expectations and outcomes as a way of working together and talking. So, there is already a healthy dialogue ongoing with business leaders and educational leaders in our jurisdiction about this issue. We really have had some useful exchanges that have helped people grasp the enormous complexity of human development and the range of capacities that young children have which need to be fostered as well as monitored.

Russell From your experience in the Toronto Board, what are the essential characteristics of benchmarks, or those which we can't afford to lose?

Sullivan First, it's really important that we preserve the detail and clarity of the benchmark system, that we keep benchmarks as models of excellence that demonstrate *clearly* to teachers, parents, and students what it is we are looking for in student work. Second, the actual nature of the tasks we are looking at needs to be emphasized, the fact that benchmark tasks are much more open ended than multiple-choice tests and promote higher order thinking skills. Benchmarks provide more than samples of student work; we've selected specific learning activities to demonstrate the qualities and levels of performance we're interested in developing in all students.

Russell Does this attention to clarity, to detail, show up in discussion at your level, Joan?

Green Absolutely. If you have a concrete illustration of what a child is doing, compared to a standard that has been set by the assessment of a whole range of children at that age, you've got something very tangible, something parents will be satisfied with, as opposed to feeling that they've been given mumbo jumbo. The concreteness of the norms we're using provides parents with an indication of our willingness to be scrutinized. I think that one of the nicest features of the Toronto benchmarks is that they are so concrete. They are not the static school-bound definitions that we had in the past but they are real, dynamic, vital and allow and encourage lively discussion between parent and teacher. When you see your child on video or performing in comparison to another, you have a different kind of experience than when you read a piece of paper.

Russell What advice, Marilyn, do you have for teachers who are tempted to begin using benchmarks in their classrooms?

Sullivan I think it's important for teachers to recognize that benchmarks are not outside the normal classroom practice, that, in fact, the models illustrated through benchmarks are drawn from normal, average classroom experiences. So, the first thing is to eliminate the fear that this type of assessment is externally imposed, or somehow different than what their curriculum has promoted for the last decade. Then teachers need to recognize that parents and students themselves are now looking for much more precise indications of the quality of student work. So, spending the time, becoming familiar with a benchmark approach and starting to internalize criteria for quality performance, is a really worthwhile enterprise.

Russell Joan, what advice would you

> We now have very little time to assure people about standards and also to provide some way in which our teachers can talk with one voice on these issues. We began to identify these as early as 1984. The luxury of time is just not available in the same way today.
>
> — Joan Green

give to CEOs who are thinking of starting up a benchmark program now?

Green I don't think there's ever been a time, at least in my memory as an educator, when we were under such enormous scrutiny. And in terms of the fiscal restraints we face, a beginning point today would have to be much more accelerated than ours was. We now have very little time to assure people about standards and also to provide some way in which our teachers can talk with one voice on these issues. We began to identify these as early as 1984. The luxury of time is just not available in the same way today. Yet the advice I would offer anyone who was looking at this initiative is based on what I believe is part of our obligation as providers of educational service to children, and that is to ensure they understand how well they're doing and that their parents are engaged meaningfully right from the beginning as well. We know from all the research that children will attain much, much more success if their parents know what it is they're doing. So I would make that connection first, between accountability and parental involvement, and make it very clear we believe it to be (a) our obligation and (b) our pleasure to work with parents because we know that in so doing we'll be more effective with the children.

Now, with respect to the communication flow from the CEO's office, I think it is most important to ensure that the policy makers understand what it is the implementators are doing. It's one thing to have a board hear the general aims of a policy and then approve it. It's another thing to have them understand it well enough that they can, among their own numbers, count on spokespersons for that policy. In our board, several trustees have been very outspoken advocates of our work, and people like Fiona Nelson are called upon regularly to speak on benchmarks in the media. So, first, you have to make sure your trustees are with you, and that is a fundamental task of the director of education.

I think the other thing the director has to do on a very regular basis is talk about the advances and accomplishments and victories that have been made and tie those to people's work so that they see in a sense a picture of cohesion. In other words, if any accountability scheme is seen as an add-on or imposition and not as something integrated into the life of the school and the priorities set for the system in terms of children's learning, then all that systemic institutional resentment begins to get in the way and emerges as yet another obstacle. So, there is a tremendous need to reinforce new policies at every turn. I did a lot of visiting of schools in my first year as director. I still teach two hours every week. And I always make a point in working with teachers to ask them to give me their latest insights on the benchmark program. That's actually how I stay on top of what's going on in schools.

Finally, I'm still of the school that believes principals are largely responsible for the work of staff and the achievement of curriculum goals. So we need to work with our principals to keep priorities straight. It is very important for the director of education to be an advocate of good program policy and practice, not just a person who balances and crunches numbers.

So, we need the people leading organizations to be involved in understanding the partnership between parents and the school in terms of accountability for children's performance. We need to keep politicians who set policy in the "loop" of understanding and we need to keep them informed about developments and also we have to be candid with them about our failures or the ground yet to be broken. One of the things I face all the time from politicians is, "These benchmarks are great theoretically but I still have 16 parents calling a week saying Johnny is not doing what Suzie's doing." So, there's a tremendous onus on all of us to make sure that that universal understanding is there and politicians have to be told there is still a struggle. We have to devote whatever fiscal resources we can to this activity, and reinforce the human resources wherever we can, and we also have to be more vigorous and outspoken in defence of curriculum and program priorities than we have been in the past because as our resources become seriously challenged there are those who would be reductionists and minimalists. That's one of my obligations: to try and keep the system focussed and thinking in a "big picture" way.

Russell Thank you for sharing your experience with us. I've enjoyed talking to you.

The first thing is to eliminate the fear that this type of assessment is externally imposed, or somehow different than what their curriculum has promoted for the last decade.

— Marilyn Sullivan

Demystifying the Goals of Education

Toronto's Benchmark Program

Sylvia Larter

Research Officer (Benchmarks),
Research and Development Centre,
Toronto Board of Education

James Donnelly

Principal, Whitney Public School,
Toronto

For more than 30 years, the City of Toronto chose not to evaluate its elementary student population against any internal or external standards. Teachers at Toronto's more than 100 elementary schools assessed students according to their own preferred methods, including self-made tests, cumulative samples of students' work, and teacher-student conferencing. For reporting to parents, they developed their own schemes — ranging in style from highly personalized anecdotal reports to numbers and computer-coded comments.

Toronto's approach to evaluation afforded teachers professional control and avoided the well-documented shortcomings of standardized tests, but it also meant that teachers had few moorings to secure their judgments about student achievement. For parents there were no guarantees that teachers would not evaluate students either too rigorously or too leniently. By the mid-1980s, parents' concerns had intensified, and in June 1987, the Toronto Board of Education passed several recommendations that launched the Benchmark Program. Included in the recommendations were three guiding principles:

1. Parents have the right to participate meaningfully in decision making about their children's education.
2. Parents have a corresponding right to know, in a timely way, how well or how poorly their children are doing in school.
3. Parents are entitled to have their children's achievements determined in relation to system-wide standards.

> Toronto defines a benchmark as information which teachers, students, and parents can refer to daily as they teach, learn, and assess achievement.

What is a Benchmark?

Now, five years later, the Toronto Board of Education has developed more than 100 Benchmarks for Language and Mathematics at Grades 3, 6, and 8 representing, respectively, the end of the Primary, Junior, and Senior Elementary Divisions. In addition, work at the Secondary Division is well under way.

Toronto defines a *benchmark* as information which teachers, students, and parents can refer to daily as they teach, learn, and assess achievement. Benchmarks may be shaped in any number of ways, depending on the needs of the users, but the standards they represent are clear to all. The philosophy behind the program is that instruction, learning, and evaluation should occur simultaneously in the classroom on a continuous basis with many authentic performance activities.

In order to achieve this, teachers use the benchmarks as models in designing classroom activities that incorporate the Ministry of Education and Toronto Board of Education objectives. Teachers then observe students as they perform these activities in the classroom and evaluate students according to the levels of performance depicted in the benchmarks.

Observation and Holistic Evaluation

In developing the Language and Mathematics Benchmarks, the Toronto Board combined observation and holistic evaluation.[1] Figure 1 shows part of a Grade 8 Language Benchmark. (The part that exists on videotape is not described.) At the beginning are clusters of loosely connected, broadly stated learning objectives drawn from the Ministry and Board guidelines. Next is a brief description of a global activity, designed by committees of classroom teachers, which operationalizes the objectives. The activity represents only one of an infinite number of open-ended tasks that could be designed and has no pre-determined answer or "correct" problem-solving procedure.

An example of a Grade 3 Mathematics Benchmark task (number 16 of 26 Benchmarks) is as follows:

Geometric Solids
Classifying Objects by Various Attributes

Students sort a set of 12 wooden geometric solids in two different ways. Where necessary, students were given an explanation of what is meant by sorting. In each case, the students were asked to explain the rule they used to sort the solids into sets.

Both the language and the mathematics activities involve the use of oral language, and the mathematics activity involves concrete materials. Both tasks also require interaction between the student and the evaluator — sometimes according to a structured format, sometimes with specially designed prompts from the evaluator, and sometimes spontaneously.[2] In short, the activities resemble authentic tasks that might take place in a regular classroom setting.

Finally, because they are performance-based, the activities require that students and their work be observed. Consequently, for all such activities, the representative samples of students are videotaped.

Holistic Scoring Criteria

The development of standards, of course, suggests the need for some form of scoring. Because the Language Benchmark task in Figure 1 and in the Mathematics Benchmark activity described above had been designed in a global fashion, and because they had never before been considered in the creation of standards, developing criteria prior to the scoring proved difficult. It was more beneficial to observe the students' performances on the videotapes and simultaneously determine the criteria. Thus, the standards came to be defined as what Toronto students "can" do, not what they "should" do. Because Toronto's student population is large and diverse in its economic, social, and cultural makeup, these standards provide sound and valuable information. For the type of benchmark illustrated here, we created criteria for five levels of performance so that a standard shows the full range of the ability of Toronto students.

> Because benchmarks support the developmental whole-language approach, teachers find that the program goes hand-in-hand with their classroom practices.

The holistic criteria:

- describe many problem-solving processes, and higher-order thinking skills;

Figure 1 *Toronto Language Arts Benchmark for Grade 8*

KEY OBJECTIVES

- Use oral language to think, learn, and communicate.
- Read a painting and demonstrate an understanding that effective reading, of whatever type, is a constant search for meaning.
- Demonstrate aesthetic appreciation.

TWO PAINTINGS — VIEWING (ART BOOK) AND RESPONDING ORALLY

This benchmark was developed to assess students' ability to obtain information from pictures and to explain, describe, and evaluate that information. Two paintings, "The Fledgling" by E. Lindner and "The Young Canadian" by D. P. Brown, were chosen from the book *High Realism in Canada*, by P. Duval.

Students who did well in the activity described and explained the significant features of the paintings and gave reasonable interpretations of them. They expressed their ideas fluently with appropriate language and vocabulary. They drew reasonable conclusions from the paintings and appreciated that their ideas and feelings had value and were worthy of expression.

Before the activity, students were asked questions about paintings and were rated for prior knowledge and experience. The evaluators then showed the first painting and asked the students to describe it as exactly as possible. After responding, the students were asked: (1) What does the painting make you think about? (2) How does it make you feel? (3) Why do you think the artist did the painting this way? The evaluators next showed the second painting to the students and repeated the same series of questions. Finally, both paintings were shown together, and the students were asked to describe their similarities, make up a title for each painting, and discuss why they preferred one painting over the other. Upon completion, students were asked to evaluate the activity in terms of interest in the paintings, enjoyment of the task, and ease of describing each painting.

The Grade 6 Benchmark, L6-2, shows how Grade 6 students performed on this same task when evaluated with identical criteria.

HOLISTIC SCORING CRITERIA USED FOR EACH PAINTING INDIVIDUALLY, BOTH PAINTINGS, AND OVERALL RESPONSE

Level 5 20%	The student describes and explains the significant features of the painting(s) and gives reasonable interpretations. The student expresses opinions about the painting(s) and often links them to personal experiences. The student is able to clarify assumptions and make inferences and can explain them to the evaluator. Views are stated clearly and concisely. The student is articulate an self-confident. Feelings are considered to be valuable and worthy of expression. The student explores ideas with the evaluator using appropriate vocabulary and conversational style. The comparison of the two paintings goes beyond the obvious and is often thoughtful and insightful.
Level 4 32%	The student describes several significant features of the painting(s) with some interpretation. Some opinions about the paintings are expressed. The student's description is accurate but not figurative. Views are expressed with some fluency, and feelings are briefly stated. The comparison of the two paintings is reasonable.
Level 3 39%	The student describes and explains one or two features of the paintings with minimal elaboration. Opinions are rarely expressed, and the description may include inaccuracies. The student may be hesitant in stating feelings and may describe only a single similarity between the paintings.
Level 2 7%	The student identifies and describes a few details. No opinions are expressed. The student has difficulty expressing feelings, making comparisons, and/or articulating ideas.
Level 1 2%	The student is unable to respond or gives a very limited response.

- are true to the whole child performing a global task;
- reflect the complexity and interconnectedness among the parts of the performance;
- include, in an integrated fashion, many affective variables such as perseverance, confidence, and willingness, and some metacognitive variables such as monitoring one's problem-solving strategies;
- and, because they were drawn from students' performances, incorporate many unanticipated, as well as anticipated, outcomes and skills.

Honouring Teacher Professionalism

Benchmarks are packaged so that the standards can be observed. The missing element of Figure 1 is the exemplar of students' performances on videotape. Other benchmarks show exemplars of students' performances in print format. Teachers do not administer benchmarks to students as they would conduct traditional standardized tests. Instead, they use benchmarks as reference materials, an approach that has resulted in a variety of centralized and school-based initiatives.

To begin, all K-8 school principals and vice-principals participated in one week of in-service about benchmark materials and philosophy as well as strategies for school-based collaborative use. A few months later, school teams of administrators and teachers participated in centrally designed in-services. Since then, school-based teams continue to collaborate in implementing the Benchmark Program. The school teams are using funds from the Toronto Board and the expertise of curriculum consultants and staff development personnel to tailor the program to their unique needs. Some entire school staffs have gone off-site and devoted professional development days to the study of benchmarks. At other locations, small groups of teachers use preparation time and lunch hours to share their understandings. Teachers also take the benchmarks home and study them on their own time. The use of benchmarks as reference material promotes a collaborative culture, professional dialogue, and individual reflection.

Students use benchmarks as models of excellent performances. Parents consider them to better understand today's complex menu of educational objectives and to make meaningful decisions about their children's education.

Promoting Shared Accountability

At Whitney Public School all teachers are collaborating to implement the program. Located in a wealthy area of Toronto, Whitney's population is 476, serving students from Junior Kindergarten through Grade 6. Even though benchmarks have been designed with Grades 3, 6, and 8 students, the range of performance for each benchmark is so wide that they can be easily used across all grades.

The teachers — organized in family groups of Grades 1/2, 3/4, and 5/6 — agree on two or three benchmarks at the most appropriate level and work towards a collective understanding of the tasks, criteria, and exemplars. They then discuss the skills students need to perform the tasks, how students learn the skills, and how to teach the skills. Teachers also look closely at the role that language plays in students' demonstration of the many higher-order thinking skills associated with each benchmark.

Because benchmarks support the developmental whole-language approach (Whitney's school philosophy), teachers find that the program goes hand-in-hand with their classroom practices. During a second phase of sharing, teachers bring to the group samples of teaching/learning activities that promote the development of benchmark skills and discuss how to integrate learning and evaluation.

Illuminating Good Performances

This year the Toronto Board of Education is involving K-8 teachers in designing a new system for recording student achievement and for communicating with parents. The goal is to have the system reflect the methods and philosophy of the Benchmark Program and, thus, create consistency across the elementary and secondary levels.

As in many of Toronto's elementary schools, teachers at Whitney Public School are experimenting with a variety of reporting forms and approaches to teacher-parent interviews. They are looking forward to presenting some of their creative ideas to the system later in the year.

All of Toronto's teachers will bring to this exercise three years of experience using benchmarks to evaluate and engage students and to inform and engage parents. Unlike externally developed and scored tests, benchmarks allow teachers, students, and parents to collaborate and remain in control of learning and evaluation. Because benchmarks can be observed, they demystify the goals of education and illuminate the nature of good performances. Students use them as models of excellent performances. Parents consider them to better understand today's complex menu of educational objectives and to make meaningful decisions about their children's education.

NOTE

1. For an in-depth discussion, see S. Larter (1991), *Benchmarks: The development of a new approach to student evaluation.* Toronto: Toronto Board of Education.

2. Because benchmarks are not to be used as tests, the scripts and scoresheets designed to administer the activities have not been made available to educators.

Reprinted from *Educational Leadership*, February 1993. Copyright by ASCD.

For further information about the Toronto Benchmark Program, contact Marilyn Sullivan, Superintendent – Curriculum, at (416) 397-3808 or Margaret Evans, School Superintendent, Toronto North Education Office at (416) 393-0599.

Adapting Benchmarks to French Immersion Education

Working with a Grade 3 Classroom

Sylvia J. Spielman

Teacher, Toronto Board of Education

In 1990/91, the Toronto Board encouraged all teachers, including French Immersion (FI) teachers, to implement the new (1989) Toronto Board Benchmarks in their program.[1] Assessment is often challenging for FI teachers because not all assessment resources are designed in French or for the use of FI students. Some FI teachers questioned whether benchmarks designed on the basis of English performances would be applicable in the second-language setting. I have found, however, that many of the FI teachers' concerns may be alleviated. My conclusion is based on my use of the adapted benchmarks with my FI students and colleagues, which I shall describe in this article.

The Grade 3 FI Classroom

I followed the procedure of the Toronto Benchmark Grade 3 Activity L3-13 "Reading a Picture Book and Writing a Story." In preparation for the Benchmark activity, the students and I wrote a collective story on chart paper based on a novel that the class had studied. The children retold the story in their own words. Each child contributed to the collective story. While I transcribed the story, I edited it for correct vocabulary and syntax in an attempt to provide a good language model. The children reread the story and suggested improvements. This was followed by a choral reading of the story by the children and teacher. When the children were satisfied with their story, I helped them put together their own criteria of what constitutes a good story.

The day following this activity, the students were asked to read and write a story for *Oscar Fête Noël*. The activity was carried out similar to Benchmark L3-13. The children were instructed to "read" the picture book and then to retell the story. They were given as much time as they wanted but were told that it was an entirely "independent" activity. I emphasized that I wanted their best story-writing technique and did a quick review of their list of characteristics of a good story which was on display. Students were reminded to reread and revise their story.

School-wide Holistic Scoring Session

Copies of this class set of Grade 3 FI writing samples were distributed to a group of six FI elementary teachers, who had just attended a workshop on general holistic scoring. The first stages of the evaluation process were done confidentially. Each teacher was instructed to quickly read the class set of stories and to choose the strongest and weakest stories. The teachers then reread and placed each story in one of five categories: excellent, very good, good, weak, or very poor. Each story was then evaluated using a number from 1 to 10. Here "10"

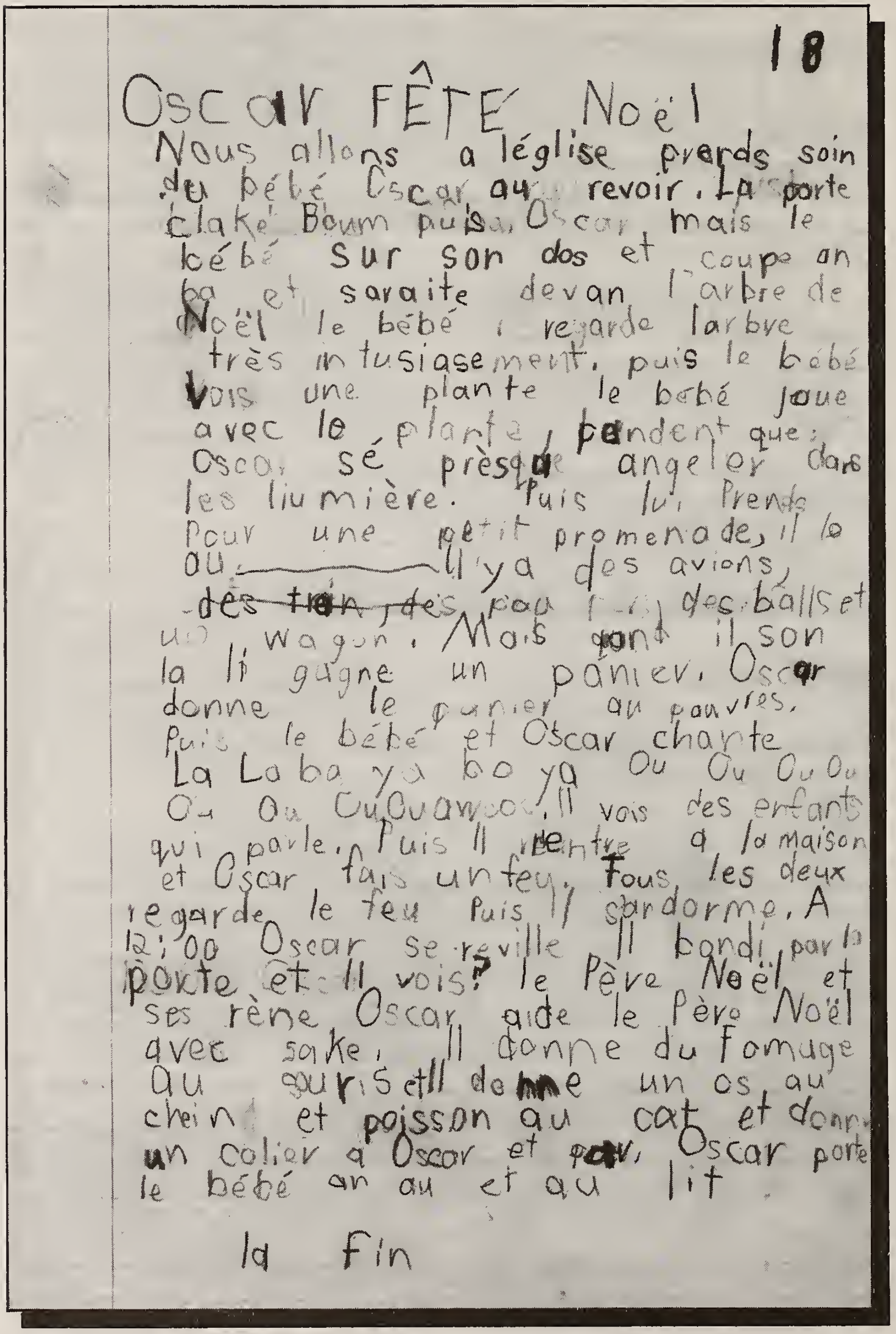
Figure 1 *Holistic Scoring Criteria of L3-13*

Taken from Toronto Board of Education Benchmarks L3-13 "Raking Leaves With Max."

LEVEL 5	The student demonstrates a knowledge of the features of stories and is aware of the needs of an audience. There is a complete telling of the story with elaboration and imaginative details that extend the information provided by the pictures. The student attempts to incorporate feelings such as happiness, annoyance, and frustration. The story is cohesive with an effective opening and conclusion. The student may attempt to use dialogue and/or interjections to highlight various parts of the story and to engage the reader's attention. The writing shows control over the conventions of written standard English.
LEVEL 4	The student demonstrates a sense of story telling and some awareness of audience. There is a complete and adequate telling of the story but with simple straightforward interpretations of the pictures. The story is fairly cohesive and has an opening and conclusion. There may be an attempt to use dialogue. The writing shows moderate control over the conventions of written standard English.
LEVEL 3	The student demonstrates some control of story form. There is an attempt to tell the story and follow the pictures. Important details of the story, however, may be omitted. Sentences may be joined by repeated use of "then" and there may be repeated use of certain phrases. The student may describe the pictures instead of creating a story. The writing shows some control of the conventions of written standard English.
LEVEL 2	The student demonstrates some control of story form. There is a short and simple telling of the story, sometimes with digressions. Some of the pictures may be skipped over. The story may have an appropriate beginning but the story telling is not sustained. The student tends to use short, incomplete repetitive sentence patterns. The writing shows some control of the conventions of written standard English.
LEVEL 1	No response or very little response.

represented the best and "1" the weakest from the entire group. Group discussion later resolved discrepancies in the teachers' ratings. After reaching a consensus, we regrouped the writings into just five levels. When we discussed the characteristics of the best — or level 5 — stories, we discovered that many of the teachers' criteria were identical to those that the children had listed. The time had now come to examine the holistic scoring criteria (HSC) developed for the Toronto Board Benchmark Activity L3-13 (see Figure 1).

Figure 2 *Narrative of Student #18*



18

Oscar FÊTE Noël
Nous allons a léglise prerds soin
du bébé Oscar au revoir. La porte
claké Boum puis, Oscar mais le
bébé Sur son dos et coupe an
ba et saraite devan l'arbre de
Noël le bébé i regarde larbre
très intusiasement. puis le bébé
Vois une plante le bébé joue
avec le plante, pendent que:
Oscar sé prèsque angeler dans
les liumière. Puis lui Prends
Pour une petit promenade, il la
ou ——— il y a des avions,
~~des trên~~, des pou [illegible] des balls et
un wagon. Mais qant il son
la li gagne un panier. Oscar
donne le panier au pauvres.
Puis le bébé et Oscar chante
La La ba ya ba ya Ou Ou Ou Ou
Ou Ou OuOuawoo. Il vois des enfants
qui parle. Puis Il rentre a la maison
et Oscar fais un feu. Tous les deux
regarde le feu Puis Il sandorme. A
12:00 Oscar se reville Il bondi par la
porte et Il vois le Père Noël et
ses rène Oscar aide le Père Noël
avec sake, Il donne du fomage
au souris et Il donne un os au
chein et poisson au cat et donn
un colier a Oscar et pav. Oscar porte
le bébé an au et au lit

la Fin

Using the Holistic Scoring Criteria

After closer examination of several samples from the class set, the teachers agreed that many of the Benchmarks' HSC were appropriate for the second-language learner. Many of the characteristics of a good story were similar in both the HSC and the teachers' list. The criteria were generic and could be adapted to assess other narratives. In addition, the criteria and their corresponding levels were developmental in nature and reflected higher-order thinking skills.

Participants were asked which criteria most needed to be modified for the second-language writings. The criterion, "The writing shows (some, moderate) control over the conventions of written standard English" was often cited as the most problematic. What is an adequate level of Grade 3, written, standard French for French Immersion students? What are the range of FI performances across the Toronto Board? What are the stages of second-language development? The conventions of written standard French of the five levels of Grade 3 French Immersion students could not be clearly defined without a lengthy analysis and discussion using FI writing samples (not included in the original Benchmark package). The FI class samples shared by the teacher groups were useful in describing the various levels of language.

An Individual Assessment

The following individual assessment shows how the benchmarks translate into the assessment and subsequent reporting to parents. Because of space limitations, the written and oral samples of only one student (Student #18) are used.

Three FI teachers in attendance at the school-wide "Benchmarks in FI" study group agreed that the criteria of level 3 best described the narrative of Student #18 (Figure 2) for the Benchmark Activity L3-13 (see Figure 1). They identified which criteria were present in the child's narrative and which were not. They compared the level 3 and 4 criteria to provide guidance in improving performance. It was noted that one of the differences in the criteria was that in level 4 "there may be an attempt to use dia-

Children's writing samples can be shared in communicating effectively with parents in report cards and interviews and among colleagues at school team meetings.

logue," whereas in level 3 "the student may describe the pictures instead of creating a story." The missing criterion of "dialogue" could become the objective of mini-lessons, (creating comic strips, role playing, etc.) to encourage this child to "interpret" (level 4) the story more through dialogue rather than "describing" (level 3) it. In addition, the teachers discussed the criterion "the writing shows some control over written standard [French]." Despite the language weaknesses identified in the text, they were also able to point out the strengths of the child's language use. Interesting words such as "Il bondi" were used (He leaped). The child risked spelling difficult words such as "intusiasement" (enthusiastically). In addition to the other benchmark criteria that characterized the narrative, they commented on the lack of organization and legibility which interfered with the comprehension of the text.

The video Benchmark L3-5 was used here to compare the child's ability to retell a story aloud as opposed to writing a narrative (L3-13). A video was made of Student #18 reading and retelling the story *Le Têtard Mystérieux*; portions were transcribed to illustrate the child's oral performance (Figure 3). It must be emphasized that the Benchmark videos were an innovative, rich, and unique contribution to the domain of assessment. This transcription loses the "life" of the retelling. The criterion, "The student retells the narrative in a story-telling mode with a sense of fun, enjoyment and confidence" is invisible to the reader as

Figure 3 *Transcript of Student #18' s Oral Performance*

L'histoire le Têtard Mystérieux est une, est une histoire avec un têtard qui ne devient pas une grenouille [child smiles] qu'il a reçue de son oncle pour son anniversaire. Et il apporte à l'école et Madame, um, et le professeur lui dit qu'elle voulait voir le têtard chaque jour pour que les enfants *puissent* [correct use of subjuctive verb!] voir comment il devient une grenouille. Mais le têtard ne devient pas une grenouille [big smile and expression in voice]! Et un jour le professeur lui dit qu'il ne veut pas qu'elle, qu'il apporte le têtard encore parce que ce n'est plus un têtard! C'est un gros têtard [says slowly and with expression]... Mais il-la-l'été passe vite et l'année scolaire recommence. Et il vient à l'école juste quand les nageurs ont plongé dans l'eau [big smile]. Aliphonse est dedans il il frappe [student hits table to emphasize "frappe"] la côté la piscine et fait l'eau toute aller sur l'entraineur [student laughs and is enjoying the mischief in the story]. Les enfants courent. L'entraineur est furieux avec Luis. Luis est le garçon. Alors l'entraineur dit qu'il doive sortir Aliphonse d'ici... Il achète le, le terrain de parking et puis il appelle quelques constructeurs, constructures et quand la piscine est finie il invite tous les enfants de la ville. La fin.

Figure 4 *Holistic Scoring Criteria of L3-5*

Taken from Toronto Board of Education Benchmarks L3-5
Miscue Analysis (Narrative) and Oral Retelling, "I want a dog."

LEVEL 5	The student's retelling accurately reflects the content and the mood of the selection. The student retells the story in sequence, links events and ideas, distinguishes between causes and effects, and makes inferences. The student monitors the organization of material. With very few prompts, personal opinions and reactions to the story and its message indicate an interpretation of the text which goes beyond mere recall. The student strongly identifies with the story. The language used frequently includes the vocabulary of the text, direct quotations (especially dialogue), or paraphrases of the text. The student retells the narrative in a story telling mode with a sense of fun, enjoyment, and confidence. Language is used effectively to recreate images of events.
LEVEL 4	The student's retelling reflects the content of the selection. The student provides sufficient summary and detail to demonstrate adequately an understanding of the story. Information is given in sequence and, with prompts, the student elaborates on details, makes inferences, and shows an understanding of the cause and effect. The student monitors the organization of material. The language is appropriate for retelling the story and may include some vocabulary from the text.
LEVEL 3	The student's retelling reflects the content of the selection. The student may provide a sparse summary. Details and events which are recalled may not be in sequence. The student requires much prompting and may make inferences to support some points in the story. Occasionally events are linked to personal experiences. There may be some monitoring of organization of material. The language tends to be less complex than that of the text.
LEVEL 2	The student's retelling contains a minimal amount of information. It is disorganized and disjointed. Irrelevant or inaccurate information may be presented. The student requires much prompting. There is little or no self-monitoring. The student tends to give one-word answers and has difficulty expressing ideas orally.
LEVEL 1	The student is unable to read the text and consequently cannot attempt the task.

is "the retelling reflects the mood of the selection." I have occasionally tried to fill these gaps in the transcription — not always very successfully.

I assessed this student's oral performance as being best characterized by the criteria in level 5 (see Figure 4). In addition, a peer-evaluation activity similar to the previous listing of characteristics of an excellent narrative was undertaken to generate a collaborative list of the peer's oral retelling performance. Again, the students' list was very similar to the HSC of the corresponding Benchmark, L3-5.

Sharing the video of this child's excellent performance with his peers seemed to boost his confidence. And building on the child's strengths in oral storytelling could help overcome writing concerns and certainly had many implications for future programming.

Sharing examples of good writing by peers proved meaningful and relevant. After samples were shared with the student, he discovered the difference between simply retelling a story (level 3 of L3-13) and creating one (as required for levels 4 and 5). The child was reminded of trying to incorporate dialogue where appropriate in subsequent writing activities as one way of working towards "creating" a story.

The classroom teacher was able to illustrate the child's oral strengths and writing weaknesses concretely to the child's parents. The print and video samples demonstrated clearly the gaps in the child's abilities. The HSC and the levels gave the parents a clearer picture of the child's present written and oral skills.

> Regular sharing and collaboration among teachers should continue to be encouraged to identify the strengths and possible adaptations of the benchmarks in the French Immersion setting.

The child's strengths and concerns were discussed and demonstrated at the parent interview. Briefly, the child's advanced oral language had to be capitalized upon to improve his writing skills. The child was encouraged to have his stories tape-recorded or transcribed by the teacher or the francophone mother. While the teacher and student worked on improving fine motor, cursive, written-language and storywriting skills, the child frequently used the class and home computer to combat temporarily illegibility, spelling mistakes, and slow rate of writing.

Conclusions

In my experience, the holistic scoring criteria provided a "stable framework as a reference point to promote discussion among teachers" (Don Rutledge, speech at OISE, March 8, 1993). In addition, use of the HSC does not prevent teachers from making any adaptations, additions, or deletions they deem necessary. Samples of authentic writing are useful and interesting to use in conjunction with the HSC. Along with the rich anecdotal information of the HSC, children's writing samples can be shared in communicating effectively with parents in report cards and interviews and among colleagues at school team meetings.

Regular sharing and collaboration among teachers should continue to be encouraged to identify the strengths and possible adaptations of the benchmarks in the French Immersion setting. Many systems of sharing are already in place: the elaborate Benchmark Libraries (Teachers' Information, Print and Video student performances), *BenchPress* (a newsletter for sharing ideas and strategies), and school and French department workshops.

French Immersion teachers need not hesitate in implementing the Toronto Board's Benchmarks in the assessment of second-language skills. Many criteria of a good FI performance are similar to the criteria derived from English performances; the emphasis on higher-order thinking skills seems to transfer across the languages. And the holistic scoring criteria, based on Ministry and Toronto Board guidelines, help explain and identify the objectives of the program. Teachers of all grades can aim towards these general objectives in planning activities suitable for their own students.

ORBIT

A subscription magazine for teachers, educational consultants, and administrators who want up-to-date professional knowledge and guidelines for practice.

Inexpensive and up-to-date. Comprehensive coverage of the latest in curriculum approaches and teaching techniques.

Four theme issues for 1993. Only $30.00 plus 7% GST for a total of $32.10

Violence in the Schools/Schooling in Violence.
Benchmarks, Standards, and Ontario Curriculum.
Progressive Approaches to Professional Development in the '90s.
Changing Schools/A Look at the Transition Years.

NOTE

1. For more detailed documentation of the Toronto Board Benchmark Program, see Larter and Spielman (1992) and Larter and Donnelly's article in this issue of *Orbit*.

REFERENCES

Day, A. (1990). *Oscar Fête Noël.* Aartselar, Belgium: Editions Chantecler.

Kellog, S. (1979). *Le têtard mystérieux.* Paris: Editions Lotus.

Larter, S. G., & Spielman, J. S. (1992). The implementation of Toronto benchmarks in the primary/junior divisions. *The International Journal of Dynamic Assessment and Instruction,* 2(2), pp. 68–76.

Toronto Board of Education. (1989). *The benchmarks: Grade three, six and eight language libraries.* Toronto: Author.

Surveying the Benchmarks

An Overview of Programs in Five Ontario Boards of Education

Christine Suurtamm

Mathematics Teacher
Peel Board of Education
Meadowvale Secondary School

A benchmark is a standard by which something can be measured or judged; it's a reference point. Let's look at the definitions of benchmarks developed recently by a number of Ontario boards of education.[1]

The Toronto Board of Education

The Toronto Board of Education began its benchmarks program because they felt it was their responsibility to parents to be able to discuss a child's achievement in relation to system-wide standards. Deciding just what those standards would be was a difficult task. The Board saw standardized tests as evaluating only a small spectrum of curriculum objectives. Also, they saw the teacher's observations as central to the evaluation of students. How, then, could standards be set?

The Toronto benchmarks are defined by what students are achieving and reflect the "normal" learning process. They are reference points for student performance which provide realistic goals and are connected to the objectives set out in the Ontario Ministry of Education's guidelines. They constitute a holistic evaluation since any one task takes into account a variety of learning objectives. Yet there are very specific criteria to guide the teacher's assessment.

The Toronto Board has recognized that it is important to discuss the use of benchmarks with classroom teachers. The benchmark library of videotapes and print material is fully accessible to teachers. The samples of student work are there to aid teachers in their own observations and evaluations, not as a set of tests for students. The samples give points of reference — or benchmarks — so that the teachers can assess where a student is *vis à vis* certain learning objectives as compared to other students in the Board.

Benchmarks are also a way of stimulating good teaching practice. Teachers often work in isolation, and using benchmarks to make classroom activities more visible and more discussable allows for enhanced professional growth. In some Toronto schools, teacher discussion groups about benchmarks have facilitated an examination of evaluation and observation practices as well as teaching practices.

The Etobicoke Board of Education

The Etobicoke Board of Education has been developing performance-based assessment techniques to ensure that assessment is cross-curricular, reflective of a variety of teaching approaches, and consistent with both Board and Provincial guidelines. These Etobicoke assessment techniques also address the broad characteristics of the learner set by the Board as objectives for students.

As in the Toronto Board, students were given tasks, or "assessment targets," which encompassed many curriculum objectives and incorporated different subjects. They were then evaluated on how well they had integrated these learning objectives. Their levels of achievement were called "indicators." This Performance Assessment Program shares with the Toronto Board of Education's Benchmark Program the goal of a holistic evaluation of the student, but is not yet as highly developed or as rigorous.

Frontenac Board of Education

The Frontenac Board of Education's benchmarks program examined levels of achievement in writing, a difficult area to assess. The Board has undertaken a System Review of Writing at various grade levels. Each year, teachers at these levels work in groups to review samples of writing from students within the Board. Using holistic scoring methods, they evaluate the students' writing samples and place them in a variety of categories.

The benchmarks in this project are the rated exemplars of students' work. The evaluated student work — the exemplars — has been catalogued and distributed to all teachers in the Board who are involved in evaluating students' writing. Each teacher then uses this catalogue of benchmarks as an aid in evaluating the work of his or her own students.

North York Board of Education

The benchmarks process in North York has addressed "learning outcomes" for reading and writing as well as for mathematics. The process has been cyclical, with constant testing, analysing, and revising. As outlined in *The Benchmark*, published by the North York Board of Education in early 1991, the project began by assessing student attitudes and achievement in board-developed standardized tests. Programs were then analysed and modified based on test results. The process involved determining measurable student "outcomes," measuring those outcomes, reviewing program effectiveness, and then developing and implementing plans for change or improvement.

Teachers were encouraged to use a variety of tools to measure these outcomes. Suggested methods included teacher observation, checklists, student journals, rating scales, self-report inventories, and tests. Teachers also used the results of the board-developed standardized tests to determine whether the students were meeting the benchmarks described in the statements of student outcomes.

The Ottawa Board of Education

The Ottawa Board of Education began the Learning Profiles Project in 1988. The Benchmarks Committee was one of four committees on this project and its purpose was to "determine the expectations for social/emotional and academic achievements at the end of each division" (Winks, 1990).

The secondary and elementary panels appeared to be taking two different approaches. At the secondary level, system-wide assessment had taken place for quite some time. A common course curriculum for all secondary level courses and benchmark "tests" used by the Ottawa Board had been completely generated by teachers in the Board. (Perhaps this concept of benchmarking is more closely related to standardization than to setting standards.)

The Learning Profiles Project was more apparent at the elementary level. Achievement targets were developed, levels (or benchmarks) of achievement for those targets were established, and examples of student work to illustrate those levels were described or collected. Most of the achievement targets were related to basic skills such as multiplication or spelling. Although achievement targets were to be assessed within the context of other work, it appeared that evaluation was focussed on those learning objectives that were easiest to measure and quantify. In comparison with the Toronto Board, Ottawa appears to be focussing on basic skills rather than evaluating the child in a more holistic manner.

> In Etobicoke, learning objectives are called "assessment targets"; in North York, they are called "learning outcomes"; and in Ottawa, they are called "achievement targets."

Benchmarking the Benchmarks

Each board is concerned with accountability in terms of the individual student's program as well as the system-wide program of education. In most boards, the emphasis has been to define what could be expected of students at various grade levels.

Another commonality is that each board has attempted to articulate a set of learning objectives and levels at which those learning objectives could be met. To some extent, each board has developed its own terminology. In Etobicoke, learning objectives are called "assessment targets"; in North York, they are called "learning outcomes"; and in Ottawa, they are called "achievement targets." The levels of achievement have been called "benchmarks" by many boards but they have also been called "indicators" or "standard reference points." The similarity in the terminology has created a similarity in structure but within that structure there are many differences.

In examining the accountability of a benchmark system, it is necessary to look at the articulation and choice of learning objectives as well as the forms of assessing whether those objectives are being met. Learning objectives range from acquiring basic skills to learning to integrate many disciplines, to being co-operative, to being a good problem-solver. Each board has taken a stance on what should be included in the assessment targets or learning outcomes. Some boards, such as Toronto and Etobicoke, have tried to incorporate a wide range of the learning outcomes set by Board and Ministry guidelines. Others seem to place an emphasis on those learning outcomes that are more easily measured by standard reference points. In assessing a board's evaluation policy, it is necessary to look at the range of learning outcomes for which a board is willing to be accountable.

Another objective for benchmark systems is the adequate assessment or evaluation of learning objectives. There is agreement among boards that student achievement *vis à vis* learning outcomes cannot be assessed solely by standardized testing. But boards also need to develop teacher expertise in evaluation. Teachers need guidance in developing assessment strategies for such learning outcomes as recognizing and using patterns, working collaboratively, and developing critical thinking. Some boards are meeting these needs better than others.

For more information on the benchmark programs in these boards, contact

Toronto Board:
Marilyn Sullivan,
Superintendent – Curriculum
(416) 397-3808

Etobicoke Board:
Anne Berens
Research and Evaluation Officer
(416) 394-7321

Frontenac Board:
John Bates
Director of Education
(613) 544-6920

North York Board:
Anne Gibson
Superintendent, Literacy Outcomes
(416) 395-8296

Ottawa Board:
Ann Jones
Assistant Director of Education
(613) 239-2211

The Toronto Board has tried to match the assessment of individual students with the broad spectrum of learning outcomes presented by the Ministry and Board guidelines (Larter, 1991). They are also attempting to help the teacher develop expertise in this type of evaluation. The Frontenac Board has used samples of holistically scored student writing to help individual teachers in their assessment of their own student's writing. It should be emphasized that a "high level" benchmark system also helps teachers grow and develop professionally in the area of evaluation.

Each benchmarking system described here is based on a model of growth. The assumption is that students will all reach the benchmark but they may reach it at different times. It is a continuum. And so it is with the boards' benchmarking systems. Some are in developmental stages while others are at the implementation phase. Just as educators don't assess a student before having an adequate amount of information, it is probably unfair to assess many of these benchmark systems at this point. But it is important to listen to the discussion and to try to make some sense of it. Through dialogue, collaboration, and looking to one another for support, each board across the province has been broadening its understanding of the evaluation issue and benefitting from being part of the discussion.

For further information, contact Chris Bridge, Exec. Superintendent, Instructional Programs (416) 890-1099.

NOTE

1. When this article was researched, the Peel Board of Education had just begun its benchmarks project, "Peel Learning Outcomes." These outcome statements and their indicators are presently in the final stage of refinement. For further information, contact Chris Bridge, Executive Superintendent, Instructional Programs (416) 890-1099.

SELECTED REFERENCES

North York Board of Education. (February, 1991). *The Benchmark*, 1(1). Toronto: North York Board of Education.

Ottawa Board of Education. (1991). *Benchmarks — Student achievement standards*. Ottawa: Ottawa Board of Education.

Suurtamm, C. A. (December, 1992). Benchmarking Ontario's benchmarks. Paper presented at Ontario Educational Research Council Conference, Toronto, ON.

Toronto Board of Education. (May 1991). *Fact sheet #6, Benchmarks*. Toronto: Toronto Board of Education.

T.V.O. (February, 1992) *Benchmarks teleconference* [video]. Toronto: T.V.O.

Winks. (1990). The learning profile project. *Learning*. Ottawa: Ottawa Board of Education.

Assessing Thinking, Understanding, and Problem Solving

Benchmarks for Grade 2 Geometry

Betty Lou Borho

Elementary Teacher,
Sir Edgar Bauer School,
Waterloo, Ontario

One of the major purposes of learning mathematics is to help individuals understand their world and to solve problems that occur in it. For children to develop sound mathematical understanding and the ability to reason and communicate mathematically, they need to be actively involved in the exploration of a wide variety of mathematical ideas. This type of exploration is dependent upon young children constructing, modifying, and integrating ideas through interacting with the real world, materials, and other learners.

Carefully designed benchmarks, as part of a much larger assessment process, can provide the means for valuing the thinking and understanding that children develop in relation to their mathematical ideas. Benchmarks can also guide teachers in working towards the right combination of activities, manipulatives, and investigations to ensure successful learning experiences.

School board curriculum guidelines are generally designed such that the broad Ministry aim becomes a more specific objective for a division or even, in some cases, for a grade level. In dealing with solid geometry, for example, the board guideline, or intended curriculum, for the Waterloo Region Roman Catholic Separate School Board contains a list of specific instructional objectives. One objective is to identify solids such as *cone*, *sphere*, *cube*, *prism*, and *pyramid*. Specific instructional objectives at the Grade 2 level include identifying and naming these solids, relating them to real-world objects, and sorting them according to their properties.

In the development of the list of board-level instructional objectives, attention was given to making the objectives clear and to emphasizing the importance of thinking in the manipulation of collectible and commercial geometric solids. Children were to have opportunities to explain and record their ideas, decisions, and experiences. In turn, teachers would be able to analyse how children think for themselves, choose manipulatives, make investigations, and find solutions to mathematical problems.

Implementing Board Guidelines at the School Level

Teachers, however, may have access to a clearly defined list of instructional objectives for a specific grade level outlined in a philosophically sound curriculum guideline. They may also have an ample supply of appropriate manipulative materials, and they may have received considerable in-service with suggestions and ideas on how to use these materials. Yet, with all of this assistance, how can we ensure that a teacher is teaching with a problem-solving focus in a specific area of curriculum? And how can a teacher be sure that the children have learned what is expected of them?

> How can a teacher be sure that the children have learned what is expected of them?

As a means of addressing these concerns, a small chunk of subject matter and the associated benchmarks can be examined closely. At our school, an "s" list (or "school" list) has been prepared for the component of solid geometry at the Grade 2 level. The "s" list is designed to clarify what is being taught in solid geometry and it contains a series of anticipated learning outcomes, appropriate and relevant to the learning capabilities of young children. The items on the "s" list have been organized so that they match with the objectives set for the topic and for maximum problem-solving experience. The "s" list should be the taught or implemented curriculum.

The "s" list for solid geometry — which may be considered a list of benchmarks — contains the following items:

- Sort solids (many groups, two groups, etc.).
- Identify individual sorting rules.
- Classify solids according to a specific rule — for example, one corner, no corners, more than one corner, and so on. (The child can identify the classification rule.)
- Select two solids that are different and explain why they are different.
- Select a *cube*, a *cylinder*, a *sphere*, and a *cone* from a collection of solids.
- Use the name of the solid (*cone*, *cube*, *cylinder*, *sphere*).
- Compare two solids (i.e., tell what is the same and different about the solids).

Given that the right balance of experiences is provided, this "s" list should help teachers identify those children who are performing at the level consistent with the benchmark criteria.

Start-Up Activities

In choosing starting-point activities, a teacher needs to decide on the type of experience necessary to identify pertinent information about the learner and to have some idea of what information can be obtained from having a child do a specific activity. Some sample activities are:

a) *Matching Solids*

Display an array of collectible and commercial solids on a table. Ask the children to:

1 Find a solid like the one in your hand.

2 Choose two solids that are the same.

3 Choose two solids that are different.

4 Have the child give some reasons for his or her choices.

5 Let the children informally play with the materials.

In these small-group activities, a teacher is able to observe and record the actions and verbal responses of the children. It is possible to get some indication of the current level of understanding for many of the children and of their ability to articulate their ideas. Benchmarks can be used to help the teacher focus on the kind of information to be collected in order to determine a starting level of performance. A teacher can gather information such as the following:

- the ability to match objects
- the rationale for matching the same objects (This may span relevant information — such as, "both can roll," "both have circles on the bottom" — to irrelevant information — such as, "both have the same colour," "both are the same size" — or no reason may be given. In some instances, a child may not be able to choose two solids that are the same or different.)
- the rationale for matching different objects (This is helpful to use with children who give an outstanding response in 2.)
- the level of previous experience with the materials (This could range from not knowing what to do with the materials to developing more sophisticated structures.)

> Benchmarks can be used to help the teacher focus on the kind of information to be collected in order to determine a starting level of performance.

Through the use of introductory activities, a teacher can notice discrepancies or gaps between anticipated responses, actual responses, and benchmark expectations.

b) *Naming Solids*

Activities such as naming solids can be selected as a means of working towards achieving a specific benchmark.

Place a set of commercial solids on the table.

1 Ask the child to choose a solid that he or she can name for you.

2 Repeat this procedure giving the child several opportunities to name other solids.

3 Ask the child which solid he or she thinks can roll the best. Which solid would not make a good roller? Which solid would not stack well?

Individual teacher-child interactions are important. It is through the use of carefully selected tasks that a teacher is able to develop a picture of what a child's performance tells about his or her level of understanding relative to the items on the "s" list. A teacher can record the number of solids that the child can correctly name. These responses can be matched against the names — *cone*, *sphere*, *cylinder*, *cube*. How many of these four solids were named by each child? What was the most frequently named solid? What other solids were named? The questions on the property of rolling can be recorded by using a tally sheet for the different solids selected. What was the most prominent response for the best roller and for the poorest roller? Is there a pattern in the responses? Often children go into long narrations about their choices and this gives information about the degree of previous experience.

c) *Describing Solids*

A game was played by a group of four children in my class. One child began by selecting a solid, hiding it behind his back, and then describing the shape for the other children, who in turn tried to guess the name of the shape. I recorded the language used in describing the solid for additional insight into the child's understanding of the geometric concepts.

Example #1
It is round.
It moves good, like straight.
It has a flat end on each side.
Is it a cylinder?

Example #2
It rolls good.
It is round.
It rolls straight.
Is it a sphere?

Example #3
It has a point at the top.
It has a flat bottom.
It moves real funny.
Is it a cone?

Example #4
It has sharp edges.
It is flat at the top and flat at the bottom.
Is it a cube?

Through examining the language used, a teacher can analyse the degree of clarity in the description, the number and nature of the clues given, the type of terminology used, and the level of sophistication of the description.

Evaluation

While there is abundant information on how to teach problem solving, there is little information on the means of obtaining data and measuring the accomplishments of children. It is difficult to determine the degree to which the attained curriculum measures up with the objectives of the intended curriculum. There are no data on how much of the intended curriculum was learned by how many children, or even which items of the intended curriculum were not taught and for what reasons.

In order to obtain this kind of data, it is essential to keep ongoing cumulative records of the children over time, as well as numbers on the amount of content that was known by the children before the teaching was started. For each item on the "s" list, a "p" value would be required to determine the "percentage of children" who learned all the items on the "s" list. Taught coverage could be obtained also for each item on the "s" list by counting the percentage of children who were taught the item.

> When dealing with benchmarks, a description of what children can do, because of what was taught to them, needs to be developed.

Benchmarks

When teachers involve children in the types of experiences required to teach the "s" list items, they are able to gather extensive information on their learning activities. Yet in order to get some understanding of what is learned and some indication of how well the intended curriculum has been taught, it is necessary to go through a process of changing the qualitative data (teacher observations, children's work samples, etc.) into quantitative data (analytical scoring). It is necessary to get scores on the learning. For example, before the teaching, only 10 percent of the children could do a specific item whereas after the teaching, 60 percent of the children could do the same item.

When dealing with benchmarks, a description of what children can do, because of what was taught to them, needs to be developed. Then, in turn, clearly defined parameters or distinguishing attributes for each level within a range of levels need to be identified for a specific set of items on the "s" list. In order to develop this set of benchmarks, or indicators, it is necessary for the teacher to examine closely some videotapes of actual classroom experiences and look at what the children are doing on the tapes relative to the specific "s" list item. The data are then analysed to determine whether the "s" list items have been taught and whether the objectives have been attained and, more precisely, by what percentage of the children.

In looking closely at small segments of classroom experiences, it is interesting to realize the extent to which perspectives change and new insights develop. Through analysing and reflecting upon the teaching/learning experiences provided for the children, a teacher can determine possible directions for change and for personal growth (see Connelly & Clandinin, 1988).

At one time I perceived benchmarks as statements of learning developed by some external source. Now, I feel that it is the teacher who is most familiar with the students and their learning interests and capabilities and who is in the best position to design benchmarks that ensure the highest level of student performance and the greatest enjoyment of the teaching/learning process.

REFERENCES

Connelly, M. F., & Clandinin, D. J. (1988). *Teachers as curriculum planners: Narratives of experience.* Toronto: OISE Press.

Mathematics Guideline K-6. (1989). Kitchener: Waterloo Region Roman Catholic Separate School Board.

Linking Opportunities and Achievement

A Program Evaluation Model for Grade 6 Mathematics

J. Stuart Telfer

Mathematics Consultant
Waterloo County Board of Education

The Review-Development-Implementation (RDI) process of implementing and assessing curriculum initiatives allows educators to think through and plan each stage of the curriculum process. However, we spend so much time fitting an educational initiative into the RDI process that it becomes of secondary importance to consider if the initiative is worthwhile doing at all, or if it will result in a measurable increase in student learning. It is essential, in my view, to re-focus our attention on what to change as well as how to change. The *what* of change must be based on analysis of the links between the learning opportunities afforded students and the measures of student achievement identified as benchmarks of performance.

I wish to explore the model that has formed the basis for ongoing evaluation of the Grade 6 mathematics curriculum in Waterloo County. The model has benchmarks that have been identified by County teachers. In fact, the strength of the model is that it invites, and indeed requires, the active involvement of teachers and mathematics leaders in the development, administration, and analysis of the test items. This process is a positive step in gaining acceptance of the evaluation model.

The model presented here is not a report card on a teacher, a school, or a system; it is, rather, a framework for identifying and directing curriculum reform. It is intended to give an accurate and ongoing series of "snapshots" of the mathematics program in Waterloo County. The data collected are meant to

- provide direction for curriculum change and teacher in-service
- provide an ongoing evaluation of the mathematics program in order to assess the effect of future curriculum initiatives
- provide information on the degree to which the intended curriculum is being taught and learned
- identify those mathematical concepts that are being taught but do not significantly improve student learning
- identify those mathematical concepts that are not being taught by a significant number of teachers but should be taught
- address the system and community needs for accountability

Outline for a Model of Evaluation

The model is based on that used in the Second International Mathematics Study (SIMS). That model looked at three curricula: the curriculum as prescribed by the system (intended), the curriculum presented by the teacher (implemented), and the curriculum received by the student (attained). The comparison of the data among the three curricula gives vital information about the degree of conformity to the prescribed curriculum and the degree to which teachers are successful in presenting the curriculum.

Once our benchmarks are operationalized, we have an effective and relevant data-gathering mechanism for defining the intended curriculum and for measuring the attained curriculum. Using the SIMS model, we can gather data about the implemented curriculum by using both a teacher and a student "Opportunity to Learn" questionnaire. Such questionnaires provide diagnostic data to detect discrepancies between the students' attained curriculum and the County's intended curriculum.

Use of this model for program evaluation requires a clear understanding of what is being evaluated and how the data will be used. Naturally, as with any type of evaluation, there will be scepticism and concern about the use of the results. Will the results be used to compare schools to schools; teachers to teachers? Teachers and administrators must be presented with a clear statement of the intended uses of the data. By emphasizing the matrix sampling techniques and by making benchmark test items freely available to teachers well in advance of the testing, teachers can be encouraged to be more receptive to the procedures and more understanding of the intent of the testing.

Large-Scale Evaluation of Grade 6 Mathematics

The Waterloo County Mathematics Department evaluated the Grade 6 mathematics program in preparation for revision to its K-6 mathematics curriculum guidelines. From September 1989 to April 1990, the Department — with the assistance of Dr. Merle Wahlstom of OISE — developed a plan for the project, created the benchmarks, and devised the data-collection mechanisms. The review focussed on:

- student achievement data for selected benchmarks in a balanced mathematics program in order to obtain measures of the attained curriculum
- student opportunity-to-learn data for selected benchmarks in a balanced mathematics program in order to obtain measures of the taught curriculum
- student and teacher attitude information related to the current mathematics program

Sample "Opportunity To Learn" (OTL) Question:

Objective D*

Which of these represents eight hundred and seven and three tenths?

A) 807.3 C) 870.3

B) 807.03 D) 873

1. During this school year, did you teach or review the mathematics concepts or skills needed to answer the item correctly?

no ... yes ...

2. If yes, please indicate in which term the major portion of this instruction occurred (select 1 only).

... Term 1, Sept.–Dec.

... Term 2, Jan.–Mar.

... Term 3, Apr.–June

3. If no, the reason it was not taught was (select 1 only):

... It will be taught later in the school year.

... Students were not developmentally ready to learn the concepts.

... Students already understood the concepts.

... The lack of time to teach the concept in the program.

... Other topics took priority.

Question #1 is used to calculate the teacher OTL index for each objective. Question #3 allows clearer study of why certain intended objectives were not taught, and Question #2 will be used to find if an increase or decrease in p-values (percentage correct) can be related to the time of year instruction is given. For example, most teachers teach Grade 6 geometry in May/June. A high p-value may be related to the close proximity of the concept to the test time rather than to improved student understanding. It may also be found that a particular topic is taught by most teachers in January/February. A significant reason for a low p-value may then be related to the time of instruction — that is, winter doldrums — rather than the difficulty of the topic.

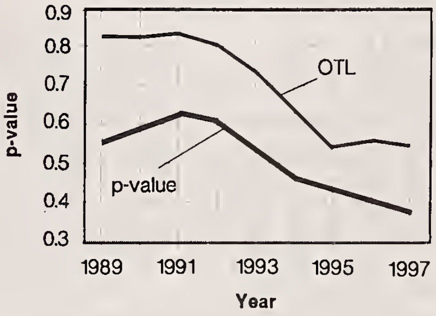
The design of the model offers both a pre-test and a post-test and time-series analysis of the data. When looking at the data from both design procedures, the central issues of causality and generalizability must be discussed. The design provides annual data relating student pre-test and post-test p-values for each objective as well as a c-value coverage index (the percentage of students taught). These data can provide valuable information to identify areas for which teacher in-service is needed.

Despite my belief that the design will produce valuable program evaluation data each year, the inherent weaknesses of a pre-test, post-test design over a one-year period require that we look at this process as part of an annual program evaluation strategy that will yield comparative data from year to year. The design offers results that are statistically sound and generalizable. The pre-test should have little or no influence on the success of the student writing the post-test and should certainly be a negligible factor from year to year. The selection process, being random, will not jeopardize the external validity. As well, since the respondents change from year to year, there will be no multiple treatment interference.

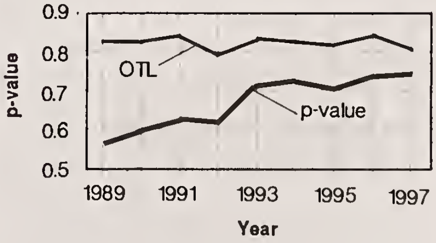
By looking at the post-test data from year to year, we can also begin to get a picture of an evolving (rather than revolving) mathematics program. I will examine two possible scenarios. By compiling annual data each year, the Waterloo mathematics community can be aware of trends in student understanding of specific mathematics benchmarks.

Objective #5: Dividing Fractions

Scenario 1



Scenario 2



In the plot shown in Scenario 1, there is a clear indication that student skill in dividing fractions is deteriorating. There is also an indication that the cause may be directly correlated to the students' opportunity to learn the skill. Is it important for students to know how to divide fractions? If, in 1997, it is not identified as an important skill, we may take the initiative to reduce the OTL further. If it is deemed important, we will need to ask whether increasing OTL will improve student performance and by how much. The data can alert educators to the situation and, possibly, to a resolution.

In the plot shown in Scenario 2, the OTL has remained fairly constant, but there is an indication of a significant p-value increase during 1992/93. What is the cause for the increase? Does this indicate the success of a particular curriculum initiative or professional development activity? Why has the p-value increased fairly consistently while the OTL has remained constant?

What these two scenarios exemplify is that on-going program evaluation may not only provide answers to many questions, but, just as important, it will also provide the rationale and the justification for these questions — questions that are based on improving student learning and identifying concepts worthy of student and teacher time to learn.

Summary

What I have attempted to describe is an active and viable process for evaluating the mathematics curriculum in Waterloo County based on the creation of benchmarks. The central theme of my article is that the mathematics community simply does not possess the data needed to direct purposefully improvements in the mathematics curriculum. We all have biased perceptions and we can all speculate, but we can no longer direct our programs based on these speculations. The model and the procedures I have described cannot and will not provide all the data necessary to direct accurately curriculum change. It is essential, however, for educators to begin the process and to begin with a model that looks at more than just student achievement scores.

The SIMS model allows an analysis of the key components in the learning process: what should be taught, what is taught, and what is learned. The model allows for the establishment of links, better learning opportunities, and benchmarks. A program evaluation designed with those parameters can provide teachers and administrators with a description of the program that is directly related to improving the classroom experience, not just to providing a scorecard on performance.

---

* These "objectives" can be considered benchmarks with a little license.

Standards, Benchmarks, and Ontario Curriculum

"How well are we doing?"

A Discussion with Bruce Archer, Burle Summers, Ken Travers, and Howard Russell

Bruce Archer
Executive Assistant,
Ontario Teachers' Federation

Burle Summers
Director, Learning Assessment Branch
and Project Manager, Provincial
Standards (Benchmarks) Project,
Ontario Ministry of Education
and Training

Ken Travers
Professor of Mathematics Education
University of Illinois at
Urbana-Champaign
formerly Director, Division for Research
Evaluation and Dissemination
National Science Foundation,
Washington, DC

Moderated by Howard Russell

Russell Ken, would you comment on "standards" and "benchmarks" as you see them from your position at the National Science Foundation?

Travers First, we're very concerned here, as others elsewhere are, with the quality of science and math that's available to our students. We use words like "world-class standards" for curriculum. And it's very helpful for us to think about issues in the international context. From my experience with the Second International Mathematics Study, where we collected data ten or so years ago, we got a lot of mileage out of distinguishing three notions, or levels, of curriculum. One was to think about what the countries expect students to know in science and mathematics — that is, the intended curriculum. Currently, we're very interested in seeing what the various educational systems expect learners to know at various age levels. And in the Third International Mathematics and Science Study, we're also looking at exit examinations around the world, for students leaving high school and going to college. What are the sets of expectations that countries have for students at that point? Yet as we all know, there's a big difference between what is in the curriculum guide (the intended curriculum) and what ends up in the classroom (the implemented or delivered curriculum). We have come up with various ways of getting a handle on the implemented curriculum. There's been a lot of creative work done on this by folks at OISE, particularly Richard Wolfe, who devised a novel way of displaying data that plots classroom data against data on the intended curriculum (opportunity to learn). These plots present dramatic evidence that classrooms differ in terms of curriculum topics that are covered. By looking at those data, on what's actually taught to students, we're better able to get a handle on the third level of curriculum — that is, student achievement.

Russell Does the word "standard," then, push benchmarks aside in your world? Or are both words going to be around for a while?

Travers We're just now getting into the notion of standards in the more empirical sense and I'm sure we'll be struggling with that for quite a while.

Russell Burle, can I ask for an update from the Ontario Ministry of Education and Training? It sounds like you're wrestling with the same problem.

Summers The question of standards has preoccupied us for many years. We've been very soft, I think generally, on how we've articulated standards. But the current demand is for us to describe them more clearly and more publicly, more universally, and to apply them more consistently. We're also struggling with an "outcomes" orientation to the question, "How well are we doing?" Ken quite rightly pointed to the fact that people expect us to go beyond our local jurisdiction, in our case the province. So the next question, once we've articulated the standard and examined the outcome against it, is "How does it compare with Japan or Germany or Britain?" There's a global aspect to the dialogue. The process of having public communication on both the outcomes or results to be achieved and the standards to be applied to them, and getting some agreement, is the current challenge.

Russell Bruce, I wanted to get your reaction to this discussion from the perspective of a teachers' organization?

Archer I back up to the concept of accountability. At one level, people just want their tax dollars to buy quality education. They want to look at a current, relevant education system. In particular, parents want to know how their kids are doing, whether they're progressing from earlier stages, relative to other kids, and meeting some understandable set of standards. It may be that the principal ingredient in all of this is to tie the public — and that includes parents and the business world — into a more open, collaborative system of education. That involvement will provide, I believe, the assurances that we are evolving standards.

Russell What do people want in terms of accountability?

Travers We tend to be intimidated when folks say that the government — that is, the people who send us money to do things — wants evidence in terms of student achievement scores that our programs are really making a difference. It's not that achievement scores are unimportant — they certainly are — but we'd like to have a lot of information. First of all, we want information about what the students are like, what the array of learning is that takes place, not only in achievement, but in attitudes and in ways of thinking about science. I mentioned earlier the notion of the delivered curriculum. We want a better handle on testing in terms of equity. One of our concerns is to what extent are the innovations that we fund promoting equity in terms of access to quality science or math education. So we like to think here that our clients are in the market for a broad array of indicators and not simply test results. Maybe part of our job is social marketing in that we need to create an interest out there in the broader view of education.

Summers Can I pick up on the notion of a "new collaboration," which Bruce alluded to earlier? As he said, we have to involve the public in education. We're hearing that from a variety of sources. But pulling that dialogue off requires sharing our turf, talking to people in an open setting where they're comfortable, not just working in the confines of a classroom. A major obstacle we have here is with the language we use. We are constantly challenged to talk more simply about these concepts so that people can participate in the debate. First, those of us in education have to collectively agree on our language and then use it consistently. Another obstacle is that the public's experience has been with scores, grades, marks, and so forth. So there's a great need for public education, as Ken suggested, about the other indicators or the pieces of evidence — and "evidence" is the word I'm using quite consistently now — which indicate achievement. What evidence would demonstrate that a student has learned? What evidence do parents have that skills have been acquired?

Travers When the results of the Second International Math Study came out, we were under a lot of pressure from the media to rank how we stood in comparison to other countries. We resisted that a great deal but were not totally successful. We insisted in our report on presenting information on sub-score level algebra and geometry and where at all possible at the item level. We had a major story on "Good Morning, America." What they did, and it was so nice I thought, was they gave an example of an item from our test, and used that to explain the standing of American kids, in the context of other countries.

Russell I think it's important, too, to be that concrete with the generation of our data. That really takes it beyond the testimonial and rhetorical.

Travers And now, with all the technology we have for making very impressive data displays, we can be much more creative in how we use the various media out there to tell our stories.

Russell Should we introduce the concept of "high-stakes testing" at this point? It does seem to have impacted on the discussion, on Burle's work at the Ministry, and I know that Bruce has thought long and hard about it because it has impacted on the lives of teachers at the classroom level.

Archer High-stakes testing is an old concern for practising professionals who worry that somehow all the evidence derived from testing will skew education, will have negative results in all sorts of ways, in terms of promotion and tenure and adverse judgments.

We all know, there's a big difference between what is in the curriculum guide and what ends up in the classroom.

— Ken Travers

Summers The high-stakes aspect of testing, which manifests itself in statements comparing students or teachers or schools, is very much a concern of the public at the moment. Even some of our politicians talk openly about their concerns for ranking students or schools. What we've tried to do is help them understand that testing for ranking isn't helpful in achieving improvement, that it isn't motivating, and it doesn't have the desired effect. It's all very well to state that, we must prove it. I believe we have to report ranges of scores and look at subscores and do item analyses, as opposed to the gross aggregrations and averages that we've tended to use in the past, and then invite people to look at very concrete areas of knowledge and skills where we have high achievement or where we need to improve. The public is smart enough to see the value of such analysis and to recognize, on a personal level, that simply ranking individuals and publishing the results either for students or schools does not have the desired effect.

The high-stakes aspect of testing, which manifests itself in statements comparing students or teachers or schools, is very much a concern of the public at the moment.

— Burle Summers

Travers I think what we really need to do is address the nature of our assessment activities. One of the most visible studies we funded here in the last while came up with the remarkable finding that our instruments emphasized low-level skills — recognition, recall, and so on — with very little emphasis on subject matter and the context of what we're trying to accomplish here at the National Science Foundation. That is, the goals for assessment were just totally counter to what we value in science and math education. So, we're putting a lot of effort now into alternative modes of assessment that emphasize the curricular goals we want to promote here. One of the major barriers right now is the testing industry and the enormous inertia not only in terms of the test companies, but also the materials in textbooks. I'm pained to say that, but the publishers have enormous control over the nature of instructional experience available to teachers and students.

Summers Just as Ken indicated, we tend to talk about very high-order goals related to problem solving, but we test at an extremely low level. That discrepancy is noted not just by educators but by the public and certainly by the business community. So, we must generate better assessment techniques. An accountant in charge of training accountants in Ontario, who had reviewed the math benchmarks, put it this way: "I wouldn't presume to tell you what to teach in Grades 3, 6, and 9, but I really like your problem orientation. We need to do that better in training our accountants so they're no longer simply bean counters but, in fact, members of a team that contributes to solving major complex problems in an organization. It's all very well to have that goal, but doing it, training people to do that, and assessing it, is the major challenge." I was pleased to get that kind of support from outside education for what we've been trying to do in assessment and evaluation.

Travers In some of the samples of assessment activities that we provided in math, we consciously avoided simple algorithms because we didn't want to give the impression that that is all there is to mathematics. We wanted to be working at a higher level. We had been told by the organizations that we were not going far enough in terms of higher-order thinking skills and problem solving. The feedback we got universally was that there was no evidence of arithmetic in what we were doing. As soon as we went back and built in arithmetic that they recognized and understood, then they were quite prepared to look at higher-order thinking skills and problem solving in mathematics that related to science and technology in a way we'd all approve. So, I guess, linking what we're trying to achieve to what people know and understand is still a critical factor in terms of implementation.

Russell Bruce, what about this direction? It seems both Ken and Burle are talking about evaluating higher-order complexity and a different mechanism for gathering the data.

Archer I think all of these comments can be used to make one point, namely that a principal reason for investigative research in testing and other forms of measurement is to improve the overall process and our understanding of it. Maybe we can never avoid invidious comparisons of teacher to teacher, board to board, jurisdiction to jurisdiction, but the right kinds of explanations about what testing can do can, I think, minimize the adverse effects. I suppose what I'm coming to is that testing ought to be constantly asking itself, "What is it we're trying to do?" And what we're not trying to do, in my view, is simply provide graphical information because that's open to misconstruction, especially in the hands of media. So, I'm very much in favour of what I'm hearing.

Russell And are you in favour of the direction that benchmarks seem to be leading us?

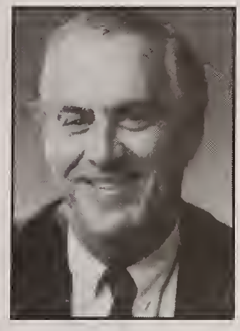
Archer Well, I think benchmarks provide a linkage between the three stages of curriculum — intended, implemented, and attained. It makes them work with each other in a mutual and complementary way.

Russell Do you see benchmarks in the same way, Ken?

Travers Yes, I do. And I reiterate the importance of the exemplars, the instructional goals we really want to achieve.

Russell And Burle?

Summers I'd like to comment by way of paraphrasing what a woman in Northern Ontario, speaking on behalf of Canada's First Nations, said to me in response to the questions about standards. Should there be standards? How should we go about setting them? How should they be used? Her comment was that standards are not the problem. It's what we do with them in schools that's



Benchmarks set a priority; they start from what students can do in a certain context, as opposed to what they should be doing.

— Bruce Archer

the problem — we continue to use them to simply fail people or to mark how deficient they are and make news stories. Frequently people know they don't meet the standard and that's why they go to school — to improve, to learn, to grow. She said that if we would use the standards to point the direction, to give clear signposts, and generate indicators of success, and then help people move along that pathway to a goal, that's what standards are for. She asked us to please get it right. I thought it was a wonderful challenge for those of us who are developing benchmarks for Ontario, what we now are calling provincial standards in mathematics and language.

Russell Can I move you now to define the most important features of benchmarks?

Summers I view benchmarks as a standards-setting process and activity. It engages the profession in being clear about their accumulated wisdom, what's generally in the heads of science or math or English teachers, making their standards explicit and inviting public response. I see it very much as a communication exercise, a dialogue with the public to gain consensus. I don't see it, as some folks do, as an end in itself, or a final act that will survive forever. I see it as evolutionary, changing, dynamic, and engaging all those affected.

Russell Is there anything emerging now in education that causes you to fear the latter possibility, that the benchmarks we have been developing will become fixed, rigid standards?

Archer I see some evidence with a minority of teachers who are not used to the public nature of this dialogue becoming defensive, retreating into a discipline and relying on internal standards that they treat as gospel. I think we need a program to dissuade the public that benchmarks are a form of standardized test. As the recent *OISE Survey of Public Attitudes to Education* showed, there is that confusion. I don't know how to avoid it but I think that is one of the difficulties. Benchmarks set a priority; they start from what students can do in a certain context, as opposed to what they should be doing. And the sequence is important. You can't really anticipate or put predictors into it until you observe what can be done in a certain context with a certain bunch of kids at a certain level. Benchmarks provide a platform, a basis for better kinds of testing, as opposed to being a test in and of themselves. I think that is one of their distinguishing — and most valuable — features.

Russell Maybe we can close by offering some advice to teachers who are concerned about the role they might play in developing and implementing assessment activities.

Travers Teachers have the day-to-day experience of working with youngsters, and so they are uniquely positioned to have a lot of input into the kinds of assessment activities that are being used. But I think it's up to folks like us to find ways to involve teachers in this process. It's far too important to leave to the professional test-makers. We allude, as we have in this conversation, to the importance of assessment activities, whatever they are, being an integral part of instruction. I think unless the teachers themselves have been involved, it's not realistic to talk about assessment as being integral. In other words, if we don't do that, we're teaching today and testing tomorrow, and never the twain shall meet. I think that's fatal in terms of any effective assessment.

Archer I think the key, and it may not be as pursuable as it was in the past, is appropriate in-service. There's lots of leadership at the board level around the province. And the Ministry is doing some yeoman service in their benchmark process which, to this point, has been an open teacher-validated process. But I go back to Ken's comment about teachers being best placed to not only judge what they're seeing and how it might be evaluated but also to recognize the need to augment what they can do at the classroom level with all sorts of assisted materials and research and strategies. I don't think that's done best individually. I think it's done collectively in an in-service context and I hope over the next few years we don't shrink the opportunities to bring teachers together around measurement and assessment issues. My advice to teachers would be to access as much structured opportunity as is available for looking at these things rather than trying to sort it all out alone.

Summers I support what Bruce said. In our experience with the development of Ontario benchmarks, we found that the most powerful part of the process was working with teachers, the kind of really excellent practitioners who reflected on their practice and generated descriptions of standards that were embedded in their thinking, in their behaviour in the classroom, and then made the description public. But the basic principle in this discussion has to be that standards apply to all. In trying to be sensitive to the needs and concerns of all students, there is a tendency, sometimes, to protect and even prevent students from knowing what the standard is or what the expectations are. I think that teachers do a super job of supporting children to grow and improve, but being clear about the standard is definitely a service to students, rather than a disservice, as long as we help them to grow and achieve.

Russell I'd like to thank you all for participating in this discussion.

What Grade for Your School?

The 1993 Orbit Essay Contest

Elizabeth Campbell

Assistant Professor
Faculty of Education
University of Toronto

When 17- and 18-year-old secondary school students were asked to assign a grade to their schools and explain their evaluations, 85 of them responded to this year's *Orbit* contest by submitting essays that were thoughtful, sensitive, humorous, and well written. The best ones captured a sense of school culture that vividly depicted life in schools across Ontario.

Congratulations to our winner — David Collins, from Pope John Paul II Secondary School in Scarborough — and to the runners-up — Jack Bennett at the University of Toronto Schools, Toronto, and Melissa Dill from Lambton Kent Composite School in Dresden. They are splitting $600 donated by OISE's Alumni Association.

It is appropriate, in this issue on Standards/Benchmarks, that evaluation is the Essay Contest theme. The judging process was itself much like "a play within a play" in which papers grading schools were, themselves, graded. While the essays were assessed on the bases of mechanics, content, and "passion" or style, the criteria used by students to grade their schools were varied.

Students used a variety of approaches in their assessments of schools. Some were largely critical. Others expressed a concern to be fair, to take seriously their evaluative task and provide a balanced, "fair" judgment. Some compared qualities across schools, while one student noted, "I've only been to this school, so I can't compare it with others."

While using varying criteria, students also reflected varying interpretations of what grades actually mean. For example, some "great" schools were assigned a B grade, and other "not-so-great" schools were also assigned a B. For educators, who also reflect similar variance in their grading, what does this imply about shifting evaluation standards across classrooms, schools, and school systems? The subjectivity of holistic grading was further evident when students from the same schools assessed the schools' merits and faults so differently that it required double-checking to confirm that they were discussing the same schools.

Despite these differences, educators may be pleased to know that most students, even given the disparity among grade interpretation, assigned their schools A's and B's; there were a few C's and one failed grade given.

Assessing Criteria: What Students Value

Most of the essays were written from the narrative first-person perspective providing an analysis of school factors; however, one graded the school through the "eyes" of a visiting friend, while another directly discussed with the school its latest report card. Another student led his readers on a conducted walking tour of his school, stopping occasionally along the way to point out the gymnasium, the science labs, the cafeteria, ...

Students provided many concrete examples to support their assessments, and common themes emerged. The academic quality of the education they receive was seen to be vital by students. Extracurricular programs in sports, music, and "special activities" were important, as was winning the district-wide school math competition. The quality of facilities for both athletic and academic pursuits was a significant factor. The aesthetic appeal of the schools was of great importance to students who preferred a pleasing, attractive "home" environment to an institutional building; community-like feelings of togetherness, friendliness, and security enhanced the good schools' ratings. Especially important to most students was the need for the school to be safe and attentive to their general well-being. One student contrasted her safe school with a "nightmare" school in which students arm themselves with knives for protection against their violent peers. In this sense, very few questioned the authority structure of the school; rules and policies were seen to enhance personal security, and most favoured them. Several students mentioned new "no tolerance" policies regarding racism, drug use, and violence, and they applauded their strict enforcement.

Of perhaps the greatest importance in evaluating schools were the interpersonal connections students make with teachers and peers. Many focussed on the exceptional teachers who make a difference in their lives on a daily basis. Others spoke of the social importance of having good friends and of the need to be aware of issues such as peer pressure, "cliquishness," and shyness.

What this year's essay contributions revealed — or, more accurately, reinforced since I expect most of us know this from our own school days as well as from our work with students — is the fact that schools are very important to students academically, physically, socially, and emotionally. Thus, striving to make all schools A+ schools is an ambitious but worthy aim.

Dr. Campbell recently won the Canadian Society for the Study of Educational Administration's Dissertation of the Year Award and was named Teacher of the Year at the Faculty of Education, University of Toronto.

ORBIT

Four
Special
Theme
Issues
for 1993
Only $32.10

Violence in the Schools/Schooling in Violence

Guest Editor: W. Gordon West, Professor of Sociology and Criminology at OISE and the University of Toronto

Standards, Benchmarks, and Ontario Curriculum

Guest Editor: Howard Russell, Professor of Curriculum, OISE

Progressive Approaches to Professional Development in the '90s

Guest Editor: Wayne Seller, Professor of Curriculum and Head of OISE's Northwestern Field Centre in Thunder Bay

Changing Schools/A Look at the Transition Years

Guest Editor: Ardra L. Cole, Assistant Professor, Department of Applied Psychology, OISE

Order your subscription NOW!

1993 rates
FOUR ISSUES
FOR $30.00
PLUS 7% GST
$32.10

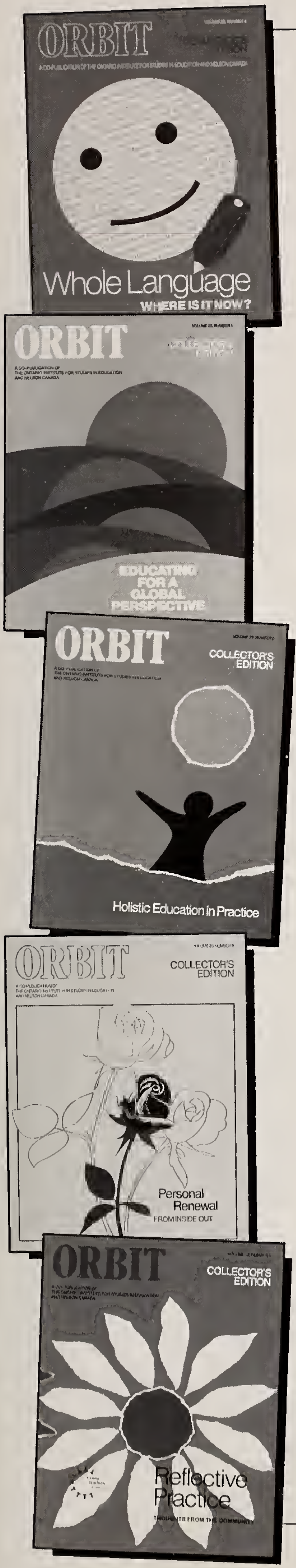
Prepayment Only

Name____
Address______________________________________
City____
Province/State_______________________ Code________
Phone_________________ Fax____________________

Mail Cheque or Money Order to:

ORBIT
P.O. Box 10 Station F
Toronto, Ontario M4Y 2L4

Phone: (416) 696-2295
Fax: (416) 696-8409



RENEW YOUR SUBSCRIPTION

Four Special Theme Issues for 1993

- **Violence in the Schools/Schooling in Violence**
  Guest Editor: W. Gordon West, Professor of Sociology and Criminology at OISE and the University of Toronto
- **Standards, Benchmarks, and Ontario Curriculum**
  Guest Editor: Howard Russell, Professor of Curriculum, OISE
- **Progressive Approaches to Professional Development in the '90s**
  Guest Editor: Wayne Seller, Professor of Curriculum and Head of OISE's Northwestern Field Centre in Thunder Bay
- **Changing Schools/A Look at the Transition Years**
  Guest Editor: Ardra L. Cole, Assistant Professor, Department of Applied Psychology, OISE

Order your subscription NOW!

1993 rates
FOUR ISSUES
FOR $30.00
PLUS 7% GST
$32.10

Prepayment Only

Name____________________
Address____________________
City____________________
Province/State____________ Code________
Phone____________ Fax____________

Mail Cheque or Money Order to:

ORBIT
P.O. Box 10 Station F
Toronto, Ontario M4Y 2L4
Phone: (416) 696-2295
Fax: (416) 696-8409

Back Issues Still Available

at $5.00 per copy plus 7% GST for a total of $5.35

☐ **Whole Language** 0030-4433
☐ **Educating for a Global Perspective** 0030-4433 01
☐ **Holistic Education in Practice** 0030-4433 02
☐ **Personal Renewal** 0030-4433 03
☐ **Reflective Practice** 0030-4433 04

Order from

NELSON CANADA, School Division
1120 Birchmount Road
Scarborough, Ontario M1K 5G4
or fax to (416) 752-9646

ORBIT

A co-publication of the Ontario Institute for Studies in Education and Nelson Canada

This is what some of our readers have to say:

The *Orbit* theme-based approach is great. Articles on different aspects of a topic give me a better understanding of the total issue.

— Marisa Cassano
Teacher-Librarian
St. Nicholas of Bari School
Metropolitan Separate School Board

Orbit is one of the few journals that blend theory and practice. The "research for the layperson" approach and the tips for teachers make it a valuable resource.

— Clare Kosnick
Instructor, Institute for Child Studies, and Doctoral Candidate,
Ontario Institute for Studies in Education

Orbit is an excellent catalyst for the professional growth of people who are committed to the successful implementation of innovative educational programs.

— Joan Green
Director, Toronto Board of Education

Most good ideas are already being tried out in one school or another but they remain unknown to the larger community of teachers and educators. *Orbit* is a must for educators interested in the latest innovative programs and ideas being put into practice.

— Michael Fullan
Dean, Faculty of Education
University of Toronto

Shouldn't you have your own subscription?

IN THIS ISSUE

The Ontario Ministry of Education and Training's Standards (Benchmarks) Project

*

Perceptions and Expectations of the Public and Business Community

*

The Toronto Board of Education's Benchmark Program

*

Other Board and Teacher Initiatives

*

ISSN 0030-4433 02